Title Applying the Finite-Difference Time-Domain to the Modelling of Large-Scale Radio Channels

Name Alvaro Valcarce Rial

Applying the finite-difference time-domain to the modelling of large-scale radio channels

Álvaro Valcarce Rial

Department of Computer Science & Technology

University of Bedfordshire

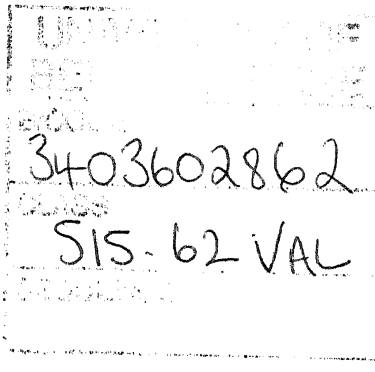
A thesis submitted to the University of Bedfordshire, in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

24 August 2010

Prediction is very difficult, especially about the future.

Niels Bohr (1885 - 1962)



Abstract

Finite-difference models have been used for nearly 40 years to solve electromagnetic problems of heterogeneous nature. Further, these techniques are well known for being computationally expensive, as well as subject to various numerical artifacts. However, little is yet understood about the errors arising in the simulation of wideband sources with the finite-difference time-domain (FDTD) method. Within this context, the focus of this thesis is on two different problems. On the one hand, the speed and accuracy of current FDTD implementations is analysed and increased. On the other hand, the distortion of numerical pulses is characterised and mitigation techniques proposed.

In addition, recent developments in general-purpose computing on graphics processing units (GPGPU) have unveiled new methods for the efficient implementation of FDTD algorithms. Therefore, this thesis proposes specific GPU-based guidelines for the implementation of the standard FDTD. Then, metaheuristics are used for the calibration of a FDTD-based narrowband simulator. Regarding the simulation of wideband sources, this thesis uses first Lagrange multipliers to characterise the extrema of the numerical group velocity. Then, the spread of numerical Gaussian pulses is characterised analytically in terms of the FDTD grid parameters.

The usefulness of the proposed solutions to the previously described problems is illustrated in this thesis using coverage and wideband predictions in large-scale scenarios. In particular, the indoor-to-outdoor radio channel in residential areas is studied. Furthermore, coverage and wideband measurements have also been used to validate the predictions.

As a result of all the above, this thesis introduces first an efficient and accurate FDTD simulator. Then, it characterises analytically the propagation of numerical pulses. Finally, the narrowband and wideband indoor-to-outdoor channels are modeled using the developed techniques.

To Yahel,
for being there all along

Contents

Acknowledgements

Doing a PhD is always hard. Specially when it is done away from home and far from your loved ones. Because of this, I would like first to thank Yahel. Although she had to walk alone the first lanes of her professional career, her continuous support and encouragement were fundamental to leading me where I stand today. Her dauntlessness, determination and constant counselling have been the best sources of advice and inspiration for both my life and work.

My parents also deserve most of my gratitude. Their typical inquiry "is a PhD helpful in finding a job?" reflects their caring and preoccupation with regard to my well-being. Despite initial worries, they have been strongly supportive the whole time and have endorsed me throughout numerous endeavours. I know they love me for what I am, regardless of the decisions I take.

My brother Sergio and my sister Laura have always shown the highest interest on my work from an outsider's perspective. It has been a pleasure, a challenge and an truly enlightening experience to explain to their bright minds what I do.

Further, it has been an honour and a joy to work in a team with so many talented researchers. Therefore, Professor Jie Zhang and the multicultural CWiND have also earned a great deal of my gratitude. The countless discussions in D109 and its small room have been the learning arenas in which the true PhD experience has occurred.

Finally, it must also be highlighted that this research was supported under the EU FP6 RANPLAN-HEC project on *Automatic 3G/4G Radio Access Network Planning and Optimisation - A High End Computing Approach* under grant number MC-EST-2005-020958.

Declaration

I declare that this thesis is my own unaided work. It is being submitted for the degree of PhD at the University of Bedfordshire.

It has not been submitted before for any degree or examination in any other University.

Name of candidate: Alvaro Valcarce Rial
Date: 8 September 2010

Signature:

List of Figures

List of Tables

List of Symbols

$\square$	Q.E.D. (quod erat demonstrandum)	
$\nabla\times$	Curl operator	11
$\nabla\cdot$	Divergence operator	11
Δ	Side length of a square grid cell	54
$\Delta\tilde{v}_{g_{aniso}}$	Anisotropy error of the numerical group velocity	88
$\Delta\tilde{v}_{g_{physical}}$	Physical group velocity error	123
ε	Electric permittivity	11
ε_r	Relative electric permittivity	67
ε_0	Electric permittivity of vacuum	67
μ	Magnetic permeability	11
μ_r	Relative magnetic permeability	67
μ_0	Magnetic permeability of vacuum	67
σ	Electrical conductivity	11
τ_{rms}	RMS delay spread	137
B	Magnetic flux intensity	11
B_c	Coherence bandwidth	140
B_D	Doppler spread	140
c	Speed of light in vacuum	67
D	Electric flux density	11
E	Electric field	11

Acronyms

ABC absorbing boundary condition

APF acceptance probability function

BER bit error rate

BS base station

CFS complex frequency shifted

CIR channel impulse response

CITI Centre of Innovation in Telecommunications and Integration of services

CPML convolutional perfectly matched layer

CPU central processing unit

CQI channel quality indicator

CUDA Compute Unified Device Architecture

CW continuous wave

CWiND Centre for Wireless Network Design

DFT discrete Fourier transform

DPSD Doppler power spectral density

DSL digital subscriber line

FAF frequency autocorrelation function

FAP femtocell access point

FDTD finite-difference time-domain

FMAD fused multiply-add

FPGA field-programmable gate-array

FRF frequency reduction factor

GPGPU general-purpose computing on graphics processing units

GPR ground penetrating radar

GPS global positioning system

GPU graphics processing unit

GSM Global System for Mobile Communications

HPBW half-power beamwidth

HSPA High Speed Packet Access

IF intermediate frequency

IHO intracell handover

INSA Institut National des Sciences Appliquées

IQR interquartile range

LOS line-of-sight

LTE Long Term Evolution

MAC medium access control

MC multipath component

MNO mobile network operator

MRFDPF multiresolution frequency domain ParFlow

NLOS non-line-of-sight

OFDM Orthogonal Frequency Division Multiplexing

OFDMA orthogonal frequency division multiple access

OPEX operational expenditure

PDP power delay profile

PEC perfect electrical conductor

PML perfectly matched layer

PSD power spectral density

QoS quality of service

RAN radio access network

RF radio frequency

RMS root mean square

RMSE root mean square error

RSSI received signal strength indication

SA simulated annealing

SIMD single instruction - multiple data

SIMT single instruction - multiple thread

STP standard conditions for temperature and pressure

SWING Smart Wireless Networking

TAF time autocorrelation function

ToA time of arrival

TVFAF time-variant frequency autocorrelation function

UMTS Universal Mobile Telecommunication System

UTD uniform theory of diffraction

UWB ultra-wideband

VSWR voltage standing wave ratio

WiMAX Wireless Interoperability for Microwave Access

WNPO wireless networks planning and optimisation

WSS wide sense stationary

WSSUS wide sense stationary with uncorrelated scattering

1

Introduction

The work presented in this thesis studies and discuses different considerations related to the utilisation of FDTD for the study of the radio channel in mobile communication systems. In particular, the indoor-to-outdoor channel in residential femtocells is considered although the techniques introduced here are also scalable to larger scenarios. Thus, this chapter presents an overview of radio propagation modelling in cellular networks with emphasis on the femtocell channel.

1.1 Wireless networks planning & optimisation

In the beginning of the twenty first century, and due to the exceptional growth of mobile telecommunications in the last two decades, radio resources are more scarce than ever before. For instance, it is common that the network infrastructures of distinct mobile network operators (MNOs) have to coexist in order to provide ubiquitous coverage and service to users. As a result, governmental spectrum regulations, investments in installation renewal, regulation of radiation limits and further economic and legislative measures have to be put in place. All of this adds up to the bills of the MNOs, who compete with each other in order to keep or increase their market shares. Nevertheless and despite the large size of the required investments, mobile telecommunications is still one of the more buoyant industries on the global stage.

Prior to the installation of a mobile cellular network in a given area, viability studies need to be performed to guarantee that the final network fulfils the requisites

for which it was first conceived. This is called the *network planning* process and it is necessary because deploying mobile cellular networks is expensive. MNOs can thus not risk the failure nor deprecation of their multimillion pounds investment. Network planning involves, among others, calculations of the required number of base stations (BSs), the achievable user data rates or estimations of quality of service (QoS). The main objective is thus the increase of revenue, which usually implies the maximisation of throughput and QoS as well as the minimisation of interference and outage. In general, the network planning process can be seen as the means of solving a multi-objective optimisation problem, for which there exist countless algorithms and approaches. It does not only involve the design and deployment of the radio access network (RAN), but also that of the core network and the links in between. In particular, the RAN can be seen as the wireless *last mile* of the connection to/from the user and it is usually the most expensive part of the network infrastructure.

The planning of wireless networks is usually linked to network optimisation, thus coining the term wireless networks planning and optimisation (WNPO). This set of techniques tries to achieve the best possible network for a given scenario. Initially, a set of constraints and network parameters, such as the minimum throughput per user or the maximum backhaul signalling, must be specified. Then, a cost function can be defined that considers the network objectives so that an optimisation algorithm can be applied to minimise it. This procedure is utilised not only prior to the commercial network deployment, but also during the lifetime of the network due to the dynamic and evolving nature of communication needs. For instance, data traffic is growing fast as a result of the increasing users demand for mobile Internet connectivity, and mobile data traffic increased 193% during 2009 in the U.S.A. only [1]. However, service revenues grew by only 29% in the same period, thus being difficult for MNOs to keep up with the pace of continuous network expansion. Therefore, the improvement and optimisation of the existing networks must take place continuously, rapidly and efficiently.

A common starting point to the wireless networks planning and optimisation process is the layout of the *candidate sites*. These are strategic locations where the MNO has already deployed BSs in the past, or where it is currently allowed (e.g. by the city hall) to install them. The decisions to be made at this stage are, among others, the final locations of the BSs as well as their radio frequency (RF)

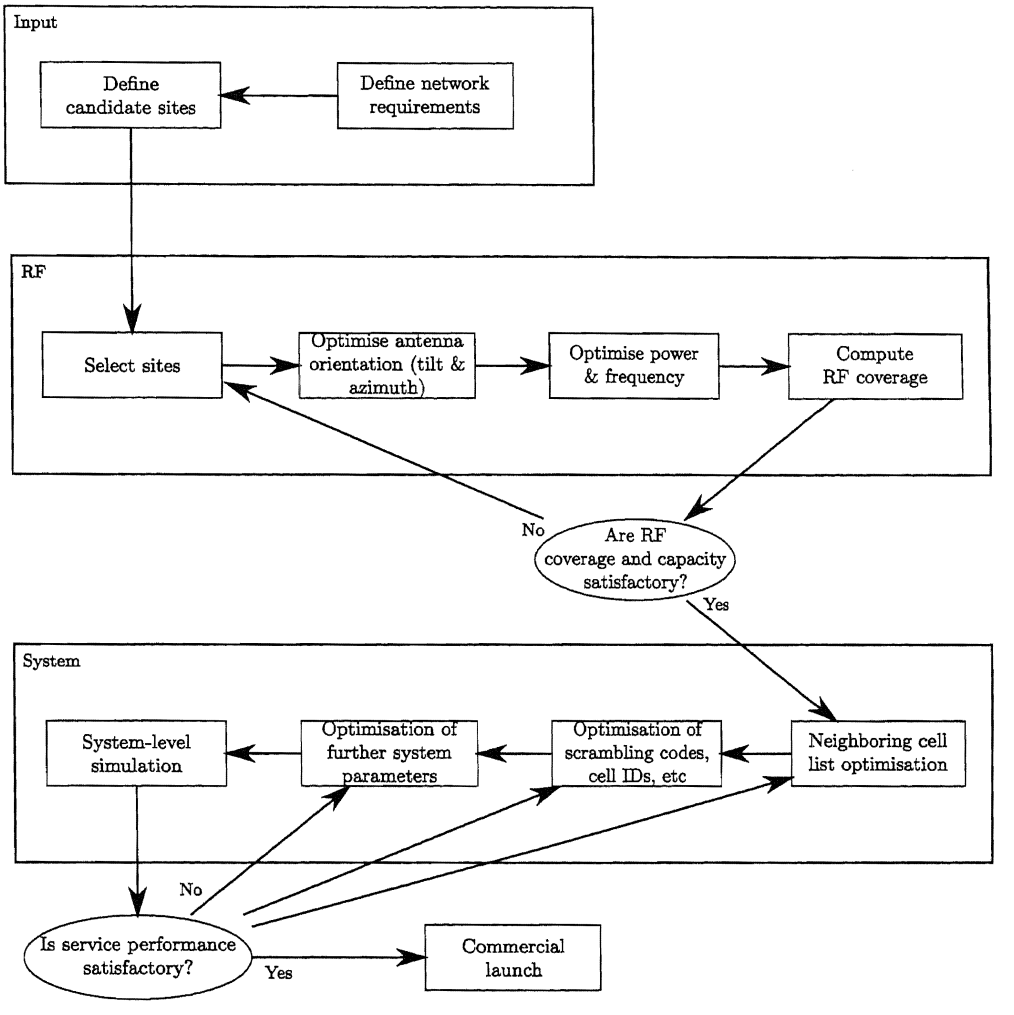


Figure 1.1: Overview of a generic pre-launch wireless networks planning and optimisation (WNPO) process from an engineering viewpoint.

settings, for which several solutions may exist (e.g. a given area may be covered by more than one BS). This can be treated as an independent problem, or together with the optimisation of other system parameters. After the RF specifications have been agreed upon, it is necessary to verify that the selected configuration suffices for the coverage and capacity requirements established beforehand. This validation is accomplished through the performance of RF coverage predictions. Moreover, it is well known that the accuracy of coverage predictions is critical for the whole WNPO process [2], as errors at this stage have long term impacts on both network

performance and capacity. The accuracy of these forecasts is usually validated via empirical measurements of radio propagation, although these are not always available. Thus, the propagation models used for this purpose must be reliable from the beginning, being their validity accepted and supported by the community.

After the RF planning has been performed, the optimisation of system parameters takes place. These vary among different technologies (e.g. GSM, UMTS or LTE), thus requiring specific approaches and techniques in each case. Then, the validation stage can be based, for instance, on system-level simulators that evaluate the performance of the network design in real-life scenarios. For instance, a dynamic system-level network simulator implementing the Wireless Interoperability for Microwave Access (WiMAX) standard could simulate mobile users in the street as they move through areas of distinct coverages. Subsequently, the achievable data rates and QoSs could be predicted based on the RF conditions (e.g. received signal strength indication (RSSI), interference, Doppler shift, etc) at a given moment. To illustrate this, Figure 1.1 shows an overview of a generic network planning and optimisation process to be considered ahead of the first network deployment. Further, post-launch optimisation is also necessary for network improvement and adaptation.

1.2 Two-tier networks

Even though mobile terminals were initially conceived as a portable communications system for outdoor usage, recent surveys [3] show that approximately 76% of mobile data access takes place indoors (home, work, etc). Consequently, the provision of indoor coverage and capacity must be thoroughly considered during the WNPO process.

During the first years of mobile cellular systems, macrocell networks were designed and deployed without much consideration to the indoor coverage problem. However as demand grew, it became clear that indoor communications were in disadvantage with respect to outdoors due to the lack of coverage. This is because the walls of houses and buildings attenuate the signal coming in from the outside, thus increasing the probabilities for outage and diminishing the RSSI and the signal quality. Then, during the first decade of the twenty first century, femtocell access

points (FAPs) were pointed out as a promissing means of, not only solving the indoor coverage problem, but also of reducing network costs and improving capacity [4]. FAPs are basically low-power base stations that have been designed for indoor usage and their objective is to allow MNOs to extend indoor coverage where it is limited or unavailable.

Through the air interface, femtocells provide radio coverage of an specific cellular standard (e.g. High Speed Packet Access (HSPA) or Long Term Evolution (LTE)), while backhaul connectivity to the core network is achieved via a broadband connection such as, for example, cable or digital subscriber line (DSL). In addition, femtocells offer other advantages to the users such as new applications, high indoor data rates, reduced indoor call costs and savings of phone battery thanks to the proximity to the BS.

Femtocells aim to provide coverage in small indoor areas such as private homes. Thus, a large area like a city can accommodate up to several thousands of this low-power base stations. The topology of a network with femtocells is hence fundamentally different from that of more traditional networks with macrocells. For this reason, a cellular network in which femtocells have been deployed is usually called a *two-tier* network, highlighting the fact that two overlapping stratums or tiers coexist geographically (the macrocell tier and the femtocell tier). With regard to the frequency assignment, two basic types of femtocell deployments can be defined:

- Co-channel: Femtocells operate in the same frequency band as the macrocells.
- Orthogonal channel: Femtocells and macrocells operate in separate frequency bands.

In addition, variations over these schemes are also possible. For instance, partial frequency reuse could also be applied, thus using the femtocells only a fraction of the whole band available to the macrocells.

In general, the deployment of femtocells has an important impact over the macrocell network. In fact, femtocells are deployed randomly from the standpoint of the MNO, i.e. it is not possible to know where the users will place their FAPs. Some may even be deployed outdoors. Thus in a co-channel frequency assignment, femtocells will most likely interfere with the macrocell network. As a result, serious

signal degradation can occur if no interference avoidance techniques are put in place. Therefore, a detailed analysis of the interference between both tiers, as well as self-configuring and self-healing algorithms for femtocell networks are necessary.

In order to cope with interference, several techniques for the optimisation of the MAC layer in two-tier networks have also been proposed. For instance, distributed spectrum allocation [5] or the use of intracell handovers (IHOs) [6] have been suggested for femtocells based on orthogonal frequency division multiple access (OFDMA). In general, these studies focus on improving performance by exploiting specific features of the network and MAC layers. Thus, most of the time the analysis of common femtocell problems with views to obtaining realistic models of the network performance requires the use of link-level as well as of system-level simulators. Similarly to the WNPO process, these simulators need to make use of RF coverage predictions as well as of more sophisticated channel models. In addition, the characterisation of radio channels is not only used by simulation-based network designs that are commercially oriented but also by theoretical and analytical studies of network problems. Therefore, radio propagation models constitute the solid pillars upon which heterogeneous wireless network problems can be studied and eventually solved.

1.3 Propagation models for mobile radio channels

Many of the network optimisation tools being used in the mobile industry today make use of empirical or semi-empirical models due to its quick implementability and short running time. Nevertheless, these models generally suffer from a lack of precission in complex environments such as dense urban or indoor scenarios, where the abundance of obstacles requires a more detailed geometrical description. In such cases, deterministic models are capable of taking more naturally the scenario geometry into account. This way, the prediction of most propagation phenomena (reflections, diffractions, refractions, transmissions and the scattering on obstacles) can be analysed in more detail throughout different locations. For this reason, most commercial network planning tools tend towards the use of deterministic models such as ray tracing, while still supporting empirical models for locations in which no topological data is available. In the following, a brief description of the two most

common types of radio propagation models for the planning of wireless networks is provided.

1.3.1 Empirical

Empirical propagation models are based on large sets of measurements data collected over one or several scenarios. In general, this corresponds to measurements of continuous wave (CW) channel power, which are later used to derive the path loss and/or narrowband fading introduced by the channel. The transmitted power, the frequency band, antenna gains, radiation patterns and further setup parameters are thus necessary to calculate the path loss from the data. If positioning information is available for all data values, then certain characteristics of the scenario can also be considered in the model. For instance, it is common to model the dependence of the path loss with the distance between transmitter and receiver. Similarly, if the geometry of the measurements site is known (e.g. terrain height, building distribution, etc), the model can be also designed to take some of these features into account. For instance, the COST-Walfisch-Ikegami model for outdoor prediction includes, among others, the street width, antenna heights and tilts in its formulation. This way, more detailed predictions can be performed than if only the distance between antennas was considered.

However, it is practically infeasible to obtain measurements in all possible scenarios. Not only it is hard to perform measurement campaigns across different sites, but also different frequencies and antenna heights need to be evaluated. Hence, the predictions provided by empirical models should not be taken as the unequivocal behaviour of the channel, but as the most probable comportment. In general, a fitting of measurements data is applied and average curves obtained. The fit residuals can then be used to model the statistical or non-deterministic behaviour of the channel. It is for this reason that empirical models are usually referred to as *non-deterministic* models, in opposition to *deterministic* ones.

The accuracy of the model is an important property that indicates how reliable the model is in scenarios akin to the one in which the measurements were obtained. First, the error of the model is defined as the difference between the model predictions and the observed measurements. Then, the accuracy can be characterised

through the root mean square error (RMSE), which is the square root of the second central moment of the error. Furthermore, it is important to evaluate each propagation model in several sites and under different setups so that its accuracy boundaries can be defined.

1.3.2 Ray-based

Ray-based models fall in the category of deterministic propagation models. In contrast to empirical models with statistical fading formulations, deterministic models always provide the same (i.e. deterministic) result for a constant set of configuration parameters. Nevertheless, they can also be used in combination with statistical models for the characterisation of random phenomena. Ray-based models derive from geometry and use basic mathematics to compute propagation trajectories. Today, there are two main ray-based techniques for radio coverage prediction. These are *ray tracing* and *ray launching*, which are summarised in the following.

Ray tracing: Ray tracing is a point to point technique that consists on finding paths between the transmitter and the receiver. The receiver location must thus be known beforehand. Rays flowing from the source are then geometrically reflected on the faces of obstacles, refracted when traversing different materials, diffracted at corners, etc. Electromagnetics-based attenuation formulas are then applied to each path to calculate the losses. For instance, when a coverage prediction is to be performed in a wide area, the site is first subdivided into cells (squares in 2D, and cubes in 3D) to create a grid. Then, ray-tracing is applied to each cell individually in order to compute the coverage prediction. The computation time is thus directly dependent on the grid size. Moreover, there are countless algorithms to trace a ray between two locations, being their complexity highly dependent on the data structure (e.g. matrices, linked lists, octrees, etc) used to store the surfaces of obstacles.

Ray Launching: Contrarily to ray tracing, ray launching is a point to multipoint technique which consists on sending off rays from the transmitter along a reduced set of angular directions. Just as in ray tracing, these rays interact with the environment, thus being subject to the common propagation phenomena. Then, the

received power at a given location is calculated as a function of the density of rays on that precise position. In this case, the number of angular increments are directly related to the computational load of the algorithm.

Most ray-based approaches assume specular reflections on every surface. However, the complex structure of modern constructions makes this assumption somewhat unrealistic. For instance, windows with double glazing and walls made of several layers of different materials do not behave physically as perfectly specular mirrors. With regard to diffractive effects, ray-based algorithms usually treat each diffraction point as a new source of rays. For instance, when a ray hits a corner or an edge, additional rays will be launched from this corner/edge and throughout a defined angular range. Although diffraction is a much more complex phenomena, this approach obtains, in general, acceptable results in the computation of the attenuation due to diffraction. In addition, the running time of these models depends highly on the detail with which the scenario geometry is defined. Hence, predictions in complex scenarios such as indoors or highly dense urban locations (e.g. London centre) are computationally expensive. Such problem is common to both ray launching and ray tracing and it is usually solved by limiting the quantity of propagation phenomena that can be computed per path. For example, each ray may be allowed to reflect up to n times, assuming that its energy after n reflections can be disregarded. The same is true for other propagation phenomena to be modelled and there is thus a limit in the precision with which these can be resolved.

As seen, each propagation model has its advantages as well as disadvantages and it is hard to find one that is fast, reliable and accurate in all possible scenarios. Furthermore, since femtocells are a fairly new technology, only a handful of propagation models exist that characterise indoor-to-outdoor scenarios at the frequencies of mobile cellular systems. For instance in 2007, the WINNER project studied the indoor-to-outdoor channel [7] via measurements in a university campus. The frequency range was between 2 and 6 GHz with an indoor antenna and at heights between 2 and 2.5 m. Then, the measurements were used to adapt the COST-Walfisch-Ikegami formula to the measured frequency range. Thanks to these, initial empirical results on the behaviour of indoor microcells were revealed. However, lower frequencies such as 900 MHz (which are of interest for femtocells based on the Global

System for Mobile Communications (GSM)) were left out of the model. Further in [8], an analytical expression for the additional losses at the indoor/outdoor interface was also developed and intended to be used with separate propagation models for the prediction of the full path loss. However, this was based on statistics extracted from ray tracing simulations and it has not been validated by measurements. According to its authors, it must thus be considered as a preliminary model.

In view of the previous works discussed above, propagation models that deal with the hybrid indoor/outdoor interface are still needed. One interesting solution lies in finite-difference algorithms, which are not commonly applied to the performance of coverage prediction in large areas. These models solve Maxwell's equations on a well-defined scenario, thus implicitely considering all interfaces between different regions. There is thus no need for separate formulations in indoors and outdoors, as the propagation of waves between environments is evolved naturally. These models are hence introduced in the following section.

1.4 Finite-difference models

Maxwell's equations, published for the first time in 1861 in his paper *On Physical Lines of Force* [9] are one of the two most important scientific achievements of the nineteenth century[1]. Today, almost 150 years after Maxwell's work was published, the catalogue of radio propagation models with application to channel modelling for wireless communications is large and heterogeneous. However, Maxwell's equations still remain as the most accurate and detailed description of how electromagnetic waves propagate. It is thus fair to say that Maxwell's equations constitute the finest radio propagation model known to mankind.

In differential form, assuming only sources of electric current and lack of free charges, Oliver Heaviside's reformulation of Maxwell's equations is

$$\text{Faraday's law of induction:} \qquad \frac{\partial B}{\partial t} = -\nabla \times E \tag{1.1}$$

$$\text{Ampère's circuital law:} \qquad \frac{\partial D}{\partial t} = \nabla \times H - J_T \tag{1.2}$$

[1] *On the Origin of Species* [10], published by Charles Darwin in 1859 is the other work disputing such honour.

$$\text{Gauss's law:} \qquad \nabla \cdot D = 0 \tag{1.3}$$

$$\text{Gauss's law for magnetism:} \qquad \nabla \cdot B = 0, \tag{1.4}$$

where B is the magnetic flux density (Wb/m^2), E is the electric field (V/m), D is the electric flux density (C/m^2) and J_T is the total electric current density (A/m^2). Also note that $\nabla\times$ is the curl operator and $\nabla\cdot$ the divergence operator. Further, the relationships between the fields are

$$D = \varepsilon E \tag{1.5}$$

$$B = \mu H, \tag{1.6}$$

with H the magnetic field (in A/m), ε the electric permittivity (in F/m) and μ the magnetic permeability (in H/m).

Most construction materials introduce isotropic, non-dispersive electric losses, thus converting part of the energy of the electric field into an electric current that dissipates heat. It is hence common to separate the total electric current density in two terms: one for the source electric current density J_s and another for the current density generated through Ohm's law. Thus,

$$J_T = J_s + \sigma E \tag{1.7}$$

with σ being the electrical conductivity (S/m), which is responsible for losses within the modelled material.

The number of practical applications of Maxwell's equations has been growing since in 1966 Kane Yee published his seminal paper *Numerical Solution of Initial Boundary Value Problems Involving Maxwell's Equations in Isotropic Media* [11]. Based on the fact that Maxwell's equations are differential, Yee introduced a numerical solution through an explicit finite-difference formulation. This technique calculates the values of the vector electric and magnetic fields on an space-time grid. Thus, in order to perform an electromagnetics prediction, the simulation scenario is first discretised (both in space and time) and a numerical source is used to excite

the grid. The simulation is then evolved by progressively computing the electric and magnetic fields, which depend on values from the previous iteration. This is known as *time-stepping*. Further, the field values from the previous iteration to be used in the computation of the current iteration are those of adjacent grid cells. This technique for spatial differentiation is known as *spatial stepping*. With this method, different materials can be easily modelled through different values of their electric permittivity, magnetic permeability and electrical conductivity. Hence, this approach is able of computing electromagnetic fields in a wide variety of complex scenarios. Finally, the technique was consolidated in 1980 on Allen Taflove's paper *Application of the Finite-Difference Time-Domain Method to Sinusoidal Steady-State Electromagnetic-Penetration Problems* [12], who coined the term FDTD. Since then, contributions from various authors have explained some of the observed behaviours, enhanced the method through different formulations and implementations, and applied it to a miscellanea of electromagnetic problems. As a consequence of nearly 45 years of research on these techniques, the current state of the art is thus a full range of finite-difference methods with different properties and applications.

1.4.1 Traditional applications

The technique introduced by Yee is easy to implement on hardware as well as on software. For instance, there are today several commercial and non-commercial FDTD solvers with applications to a wide range of subject areas. To illustrate this, some common applications of FDTD (as of 2010) are:

- Ground penetrating radar (GPR): FDTD can be used to simulate electromagnetic waves penetrating through different materials which are then reflected back to the transmitter. This procedure is then used for the calibration of ground penetrating radar (GPR) systems [13].

- Photonics: Since FDTD is a time-domain technique able of computing arbitrarily large frequency ranges, it also has application to optical and infrared frequency bands. For instance, the distribution of the electric and magnetic fields in crystals can be studied with FDTD for the design of photonic switches [14].

- Antenna design: Due to their reduced size, the computational requirements for the simulation of antennas are easily met today by most personal computers. Thus, the design of antennas is an important field of application of finite-difference methods [15] and one of the most commercially successful.

- Microwave circuit design: Microwave circuits can also be described through computational domains of reduced size. Therefore, various CAD tools exist today [16] for their simulation, test and design [17].

1.4.2 Attempts of application to radio channel modelling

As seen before, radio propagation models like ray tracing have been used during nearly twenty years. They have shown to be accurate and computationally efficient, except perhaps in environments like indoors where too many reflections need to be computed. Since FDTD is an electromagnetics model itself, it can also be considered as a feasible alternative for the purpose or radio channel modelling. FDTD is attractive because all the propagation phenomena are implicitly taken into account throughout its formulation. As a result, the last decade has seen several uses of FDTD in this area, though mainly for the purpose of radio coverage prediction.

For instance, in [18] a hybridisation of FDTD with a geometric model was proposed. In such an approach, FDTD is applied only in small complex areas and then combined with ray tracing for the more open space regions. Yet, the running time of this technique is too large to consider it for the daily planning and optimisation of wireless networks. A similar hybrid technique using uniform theory of diffraction (UTD) instead of ray tracing was proposed in [19], in which results for bit error rate (BER) prediction in one room were presented.

Further, a matrix formulation was suggested in [20] as a means for reducing the memory requirements of the standard FDTD. Then, predictions of power delay profiles (PDPs) in an urban environment were shown for a parallel implementation on a large computer cluster and at carrier frequencies below 950 MHz. This illustrated the fact that FDTD is an useful tool also for proof of concept and design validation in large scenarios, as long as efficient implementations are used. However, the execution times of this implementation were still too far from those required by WNPO techniques.

All in all, the method introduced by Yee is accurate and simple, yet subject to limitations with regard to its implementation. This is the main reason why it is not common to see commercial FDTD tools for coverage prediction nor for wideband channel modelling in large scenarios. Thus, the main challenges facing the application of FDTD to the modelling of radio channels can be summarised in the following:

- **FDTD is a computationally demanding algorithm:** A total of three (six) field vector components must be computed in 2D (3D) simulations. The computational domain is discretised in a grid and thus each of these components is stored as a large matrix, giving rise to large memory requirements. In addition, the spatial sampling step must be several times smaller than the physical wavelength, thus requiring intense numerical computations for high carrier frequencies.

- **Numerical results depend on the validity of the scenario model:** Even though FDTD approximates well Maxwell's equations, this is done for a given problem description. Therefore in order to perform an FDTD computation, a numerical model of the scenario must be obtained first. The accuracy of the prediction is thus usually as good as the scenario description. This is also true for other propagation models, although the treatment of radio propagation phenomena is more naturally characterised in FDTD. Furthermore real scenarios are always subject to random phenomena such as the opening of windows, cars passing by, random construction materials or even the movement of tree leaves. These phenomena can not be built into deterministic scenarios. Therefore, additional procedures, such as calibration must usually be considered to overcome uncertainties in the scenario description.

- **FDTD grids are subject to non-physical dispersion artifacts:** Gridding the simulation scenario is not done at a free cost. This appears in the form of numerical dispersion, which is the dependence of the numerical wave vector $\overrightarrow{k}$ with the frequency ω. Dispersion is a common phenomenon in the real world, but the physical relationship between $\overrightarrow{k}$ and ω is different from that in the FDTD domain. As a consequence of this, the propagation of numerical pulses

is subject to uncertainties, thus limiting the accuracy of wideband simulations unless high grid resolutions are used.

1.5 Focus of this thesis

The previous section has introduced some of the problems that deter FDTD from being more used by the community of RF engineers that deal with radio propagation models for large scenarios. Those engineers get usually very excited when they learn that FDTD predictions can model implicitely all propagation phenomena. The fact that, in contrast to ray-based methods, the number of reflections and diffractions in FDTD does not have to be limited, attracts also the attention of many. However, as soon as they realise that, due to the need for large grid resolutions, such predictions require supercomputers or computer clusters, and that they last for days or even weeks, the interest for this technique decays quickly. The analysis and quantification of these impediments is hence the main focus of this thesis, being the specific targets described in the following.

1.5.1 Aims & objectives

The planning and optimisation of wireless networks is a dynamic process that takes place during the whole life of the network and on a daily basis. In order to test and analyse different solutions to a given network configuration problem, computer simulations at the link and system levels have to be performed. Furthermore, it is common that hundreds of different solutions may exist to a certain problem, thus leading to different network performances and solution costs. For example, in order to provide a minium service in a given area, potential solutions could involve installing new antennas or even reorienting some of the existing ones. The validation of such solutions would require the performance of several new coverage predictions. Since decisions must be taken fast to avoid high operational expenditure (OPEX), propagation models with calculation times on the order of days are simply not acceptable. Thus, in order for FDTD to become a common and feasible alternative to other more popular approaches, speed in the computation is a decisive factor for most RF engineers. To overcome this barrier, part of the work presented in this

dissertation advances the subject area of propagation modelling by introducing a new implementation that enhances the speed of FDTD simulations while guaranteeing the accuracy in the predictions. These contributions are presented in chapters 3 and 4.

Attenuation is one of the most important signal degradations introduced by the wireless channel, since it determines the power level collected by the receiver. As a consequence of this, the necessary signal amplification is defined, thus having also an impact on the whole link budget. The prediction of path loss is hence the main purpose of radio coverage predictions obtained throughout radio propagation models. However, the wireless channel introduces other distortions that are not fully described just by the path loss. For instance, the in-phase and out-of-phase addition of multipath components (MCs) that reach the receiver with slightly different delays gives rise to fading. Furthermore, the time separation between the arrival of the first MC and the last MC imposes a limit to the speed with which pulses can be transmitted through the channel. This type of wideband characterisation is of outmost importance in several communication systems such as, for instance, ultra-wideband (UWB).

The time-domain formulation of FDTD is thus an attractive feature to be exploited for the performance of wideband studies, since arbitrary waveforms can be modelled. However, the non-physical dispersive properties of FDTD are an important source of error in wideband simulations. The effects of numerical dispersion are insignificant when the grid is discretised using small spatial and time steps. Though, the computational requirements grow with the grid resolution. Therefore, the quantification of the non-physical error introduced in the simulation by low resolutions must be considered. If this error could be predicted analytically, then custom FDTD grids could be designed so that the error bounds were known for a given set of computational requirements. However, even though some work has been already done (e.g. [21] or [22]) on the simulation of pulsed sources with FDTD, none of these studies uses low spatial and time resolutions, thus being the computation times usually far too long (days or even weeks).

Due to all of the above, the most significant contributions of this dissertation are related to wideband FDTD simulations. These are presented in chapter 5 and comprise two analytical models of the behaviour of numerical wavepackets in FDTD

grids. More specifically, these models describe in detail the uncertainties introduced by the grid in the propagation of numerical pulses: the anisotropy of the numerical group velocity, and the time spread of numerical pulses. Thanks to these results, future users of the FDTD method can, from now on, configure their simulations to achieve a desired level of accuracy, while minimising the computational costs.

To summarise all of the above, the main aims of this dissertation are:

- To characterise theoretically the error due to numerical dispersion in wideband FDTD simulations.
- To establish a methodical procedure for the evaluation of wideband FDTD simulations.
- To design and build an FDTD simulator which is fast enough for studies of large-scale radio channels, including:
  - Narrowband coverage predictions for the purpose of WNPO
  - Wideband channel impulse response (CIR) predictions

1.5.2 Motivation & thesis organisation

The implementation and calibration techniques of **chapters 3 and 4**, as well as the wideband studies of **chapters 5 and 6** are general, and can thus be applied to electromagnetic problems of arbitrary size. However, I did not obtain my inspiration to research these very specific FDTD problems from abstract discussions on the general applicability of the model, but from the specific objectives determined by the project that supported my research. It must then be indicated that the work presented in this thesis has been funded by the EU FP6 RANPLAN-HEC project on *Automatic 3G/4G Radio Access Network Planning and Optimisation - A High End Computing Approach* under grant number MC-EST-2005-020958. The main objective being seeked was from the outset the obtainment of novel planning and optimisation methods for wireless networks, both at the RF as well as at the system level. With this target in mind, my PhD advisor put together a team of researchers to work on tasks such as propagation modelling, network planning, simulation and optimisation.

Since I had experience in the design of empirical radio propagation models [23], I started to look into innovative approaches to coverage prediction for WNPO. That is how, when I started my PhD in 2007, I met Dr. Guillaume De La Roche who was writing his PhD thesis [24] about the multiresolution frequency domain ParFlow (MRFDPF) [25]. This is a frequency reformulation of ParFlow, a method based on lattice Boltzmann techniques that had been previously applied to the performance of radio coverage predictions [26]. Because of computational reasons, ray-based models are well known for limiting the amount of reflective and diffractive phenomena that can be accounted for. Thus, I found promissing the idea of using finite-difference methods for coverage prediction, as propagation phenomena were implicitely considered by the formulation. Further, I learned that there had been attempts to use FDTD for radio coverage prediction as early as of 1993 [27] in very small indoor rooms. In addition, some CIR predictions were even done in 1995 [28], although within the computational limits of that time and obviating the errors introduced by the grid.

Since FDTD is a natural parallel algorithm, a parallel implementation seemed a reasonable approach. Fortunately, when I started my PhD in 2007, *NVIDIA* also released the Compute Unified Device Architecture (CUDA), which is a set of *C* libraries that interface with the computing power of graphics processing units (GPUs). With this, the necessary tools for a parallel implementation were only a computer with a graphics card and knowledge of the *C* language. After some quick calculations, it soon became obvious that the memory requirements for 3D simulations were too high even for some GPU clusters. Thus, efforts were put on designing an efficient 2D implementation that could later be calibrated using empirical measurements of radio propagation. This is described in detail in **chapter 3**.

Furthermore, around 2007 femtocells started to become a realistic solution to the indoor mobile coverage problem and several research challenges on this topic arose. The coverage area of a femtocell is not limited to the house that hosts it. Thus, *open access* was proposed as a solution to the interference problem [29], hence creating a scenario in which outdoor users need to connect to the indoor FAP. This scenario was not common before the appearance of femtocells and there was thus a lack of radio propagation models for such a channel. Since FDTD is formulated coherently

and without distinction between the indoors and the outdoors, its application to femtocell scenarios is straightforward. Furthermore, my research group, the Centre for Wireless Network Design (CWiND), began to work intensively on femtocell related problems and hence, the need for accurate and fast coverage predictions to do these studies was imperative. Therefore, **chapter 4** describes a novel FDTD calibration method, as well as coverage predictions in femtocell scenarios.

In addition, next generation cellular technologies such as LTE will use large frequency bands (up to 100 MHz in LTE-Advanced). Further, the LTE standard introduces, for the first time, support for home eNodeBs (femtocells). This reinforces the need for wideband channel models in indoor-to-outdoor scenarios. Hence, the applicability of the work on FDTD presented here is exemplified with femtocells. In particular and without loss of generality, residential indoor-to-outdoor sites are used throughout this thesis to demonstrate the use and performance of the developed techniques. For the reasons expressed above, **chapter 5** introduces a theoretical framework for the configuration of wideband FDTD simulations. Then, **chapter 6** uses the developed theory for the wideband characterisation of the femtocell radio channel.

Finally and to ease the understanding of the formulation and methodology used throughout this thesis, the following chapter introduces and describes the employed mathematical and hardware tools. **Chapter 2** is thus intended to be read as a summarised tutorial on the fundamentals of FDTD, GPU computing and electromagnetic measurements equipment.

2

Tools & methods

Taking Maxwell's equations as an starting point, section 2.1 in this chapter introduces the theoretical background of the finite-difference time-domain (FDTD) method for electromagnetics computation. First, the 2D discrete field update equations are presented, and then, the main algorithm parameters affecting the dispersive properties of the grid are explained. Abundant literature already exists about this topic (e.g. [30] or [31]). Nevertheless, the concepts described here are intended to support the understanding and comprehension of the research results presented in chapters 4 and 6, thus making this thesis self-contained. Hence, note that the theory found in the following is summarised with respect to what can be found in specialised texts.

Thereafter, additional tools used during the development of this thesis are also described. In section 2.2, the perfectly matched layer (PML) for the absorption of outgoing waves is introduced. The use of this technique is necessary to avoid that reflections on the border of the computational domain introduce errors during the simulation of unbounded regions. Further, the necessary equations for an efficient implementation of the PML called the convolutional perfectly matched layer (CPML) are also presented.

Then in section 2.3, the Compute Unified Device Architecture (CUDA), which is the computing platform used for implementing the FDTD algorithm, is described. The parallel processing flow, as well as the different memory access schemes of CUDA are explained here. These concepts will be then used in section 3.3 to describe the implementation guidelines proposed in this thesis.

Finally, the hardware equipment used for the recording of coverage measurements is presented in section 2.4. This includes descriptions of the antennas, the signal generator and an spectrum analyser.

2.1 The finite-difference time-domain method

The electric and magnetic field are vector magnitudes. Let us then in the following, denote the vector electric field by $E = (E_x, E_y, E_z)$ and the vector magnetic field by $H = (H_x, H_y, H_z)$. When the structure to be modelled is constant along an arbitrary dimension $\zeta \in \{x, y, z\}$, i.e. it extends to infinity along ζ, the electromagnetic formulation can then be reduced to one of two possible two-dimensional cases. There is thus a *transverse magnetic mode* (TM_ζ) and a *transverse electric mode* (TE_ζ), depending on the field that is orthogonal to ζ. This can be used to approximate, for instance, a residential area, in which the walls of the houses along the street are large along the vertical dimension from the viewpoint of azimuthal wave propagation. Furthermore, since in the specific case of femtocells, the transmitter is indoors, rooftop-to-street propagation is negligible, being reflections, transmissions and diffractions through doors and windows the main propagation phenomena. Next, let z denote the vertical dimension and consider the TM_z mode (akin to vertically polarised fields) and Maxwell's equations, which were introduced already in equations (1.1) to (1.4). The TM_z formulation in Cartesian coordinates is then

$$\begin{aligned} \frac{\partial H_x}{\partial t} &= -\frac{1}{\mu} \cdot \frac{\partial E_z}{\partial y} \\ \frac{\partial H_y}{\partial t} &= \frac{1}{\mu} \cdot \frac{\partial E_z}{\partial x} \\ \frac{\partial E_z}{\partial t} &= \frac{1}{\varepsilon} \cdot \left[\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} - (J_s + \sigma E_z) \right]. \end{aligned} \tag{2.1}$$

What Kane Yee proposed in his seminal paper [11] of 1966 was to calculate the partial derivatives of Maxwell's equations using a finite-difference approximation which is second-order accurate in space and time. To illustrate this, let us in the following denote the discrete coordinates of an arbitrary point in a 2D spatial grid by (i, j), which corresponds to the physical coordinates $(i\Delta x, j\Delta y)$, with Δx and Δy

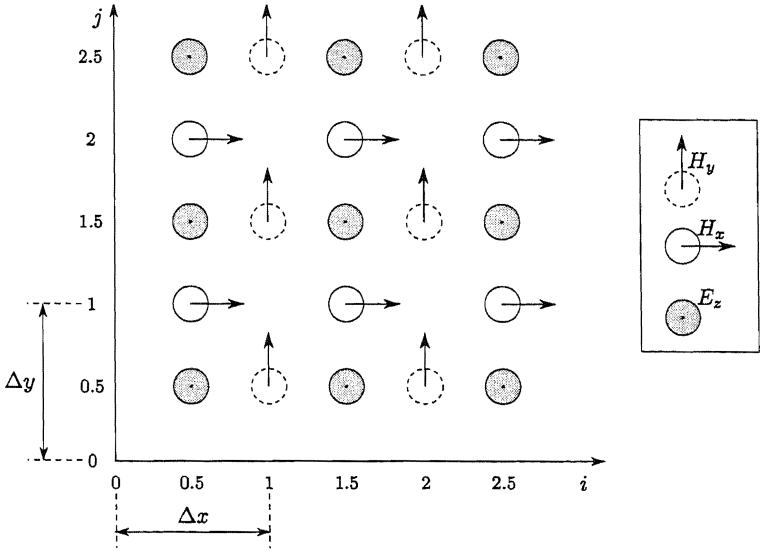


Figure 2.1: Staggered Yee grid for the TM_z mode.

the spatial steps in the x and y directions. Similarly, let n denote the discrete time step corresponding to the time moment $n\Delta t$ with Δt the time increment. Then, an arbitrary magnitude f, which is a function of space and time (e.g. the electric or the magnetic field) and is evaluated at $(i\Delta x, j\Delta y, n\Delta t)$ in a 2D discrete space-time grid is to be denoted in the following by $f|^n_{i,j}$.

Furthermore, Yee disposed the field components in an interleaved manner throughout the grid. For instance in this case, if a given H component is located at coordinate i, then two adjacent E cells will be located at $i-\Delta x/2$ and $i+\Delta x/2$ respectively. This spatial arrangement is illustrated on Figure 2.1 and is called an *staggered* Yee grid. This approach allows the second-order approximation of the spatial derivative of E to be computed at the location of H. In addition, the E and H fields are also interleaved in time in such a way that the H field is calculated at time n, then E at time $n+0.5$, then again H at $n+1$, E at $n+1.5$, etc. This time layout is called a *leap-frog* discretisation.

According to this arrangement, the equations of (2.1) are approximated as follows:

$$\frac{H_x|^{n+1}_{i-0.5,j+1} - H_x|^n_{i-0.5,j+1}}{\Delta t} = -\frac{1}{\tilde{\mu}_{i-0.5,j+1}} \cdot \frac{E_z|^{n+0.5}_{i-0.5,j+1.5} - E_z|^{n+0.5}_{i-0.5,j+0.5}}{\Delta y} \tag{2.2}$$

$$\frac{H_y|_{i,j+0.5}^{n+1} - H_y|_{i,j+0.5}^{n}}{\Delta t} = \frac{1}{\tilde{\mu}_{i,j+0.5}} \cdot \frac{E_z|_{i+0.5,j+0.5}^{n+0.5} - E_z|_{i-0.5,j+0.5}^{n+0.5}}{\Delta x} \tag{2.3}$$

$$\frac{E_z|_{i-0.5,j+0.5}^{n+0.5} - E_z|_{i-0.5,j+0.5}^{n-0.5}}{\Delta t} = \frac{1}{\tilde{\varepsilon}_{i-0.5,j+0.5}} \cdot \left[\frac{H_y|_{i,j+0.5}^{n} - H_y|_{i-1,j+0.5}^{n}}{\Delta x} - \frac{H_x|_{i-0.5,j+1}^{n} - H_x|_{i-0.5,j}^{n}}{\Delta y} - (J_s|_{i-0.5,j+0.5}^{n} + \tilde{\sigma}_{i-0.5,j+0.5} E_z|_{i-0.5,j+0.5}^{n}) \right]. \tag{2.4}$$

Note that, in these equations, the tilde ˜ identifies variables with significance only in the numerical domain. For instance ε, μ and σ represent physical properties of materials (electrical conductivity, magnetic permeability and electrical conductivity), while $\tilde{\varepsilon}$, $\tilde{\mu}$ and $\tilde{\sigma}$ represent their numerical equivalents which do not, in general, have the same values.

Also, it is observed from equation (2.4), that the term $E_z|_{i-0.5,j+0.5}^{n}$ on the right hand side has not yet been evaluated at the time of computing the left hand side term $E_z|_{i-0.5,j+0.5}^{n+0.5}$. This is because only the value of E_z from the previous time step ($E_z|_{i-0.5,j+0.5}^{n-0.5}$) is known. This equation can thus not be solved directly. Nevertheless, this term can be well approximated using a semi-implicit approximation [31], which consists on calculating it through its temporal average:

$$E_z|_{i-0.5,j+0.5}^{n} = \frac{E_z|_{i-0.5,j+0.5}^{n+0.5} + E_z|_{i-0.5,j+0.5}^{n-0.5}}{2}. \tag{2.5}$$

Rearranging now the terms and following the notation of Taflove in [31], the fully explicit formulation of the FDTD equations for the TM_z mode is

$$H_x|_{i-0.5,j+1}^{n+1} = H_x|_{i-0.5,j+1}^{n} - D_b|_{i-0.5,j+1} \cdot \frac{E_z|_{i-0.5,j+1.5}^{n+0.5} - E_z|_{i-0.5,j+0.5}^{n+0.5}}{\Delta y} \tag{2.6}$$

$$H_y|_{i,j+0.5}^{n+1} = H_y|_{i,j+0.5}^{n} + D_b|_{i,j+0.5} \cdot \frac{E_z|_{i+0.5,j+0.5}^{n+0.5} - E_z|_{i-0.5,j+0.5}^{n+0.5}}{\Delta x} \tag{2.7}$$

$$E_z|_{i-0.5,j+0.5}^{n+0.5} = C_a|_{i-0.5,j+0.5} \cdot E_z|_{i-0.5,j+0.5}^{n-0.5} + C_b|_{i-0.5,j+0.5} \cdot \left[\frac{H_y|_{i,j+0.5}^{n} - H_y|_{i-1,j+0.5}^{n}}{\Delta x} - \frac{H_x|_{i-0.5,j+1}^{n} - H_x|_{i-0.5,j}^{n}}{\Delta y} - J_s|_{i-0.5,j+0.5}^{n} \right] \tag{2.8}$$

where D_b, C_a, and C_b are the update coefficients that depend on the properties of the different numerical materials:

$$D_b|_{i,j} = \frac{\Delta t}{\tilde{\mu}_{i,j}} \tag{2.9}$$

$$C_a|_{i,j} = \frac{1 - \frac{\Delta t \cdot \tilde{\sigma}_{i,j}}{2 \cdot \tilde{\varepsilon}_{i,j}}}{1 + \frac{\Delta t \cdot \tilde{\sigma}_{i,j}}{2 \cdot \tilde{\varepsilon}_{i,j}}} \tag{2.10}$$

$$C_b|_{i,j} = \frac{\frac{\Delta t}{\tilde{\varepsilon}_{i,j}}}{1 + \frac{\Delta t \cdot \tilde{\sigma}_{i,j}}{2 \cdot \tilde{\varepsilon}_{i,j}}}. \tag{2.11}$$

It is seen from these equations that the fields at each iteration are calculated from the values they had on the previous iteration. Thus, starting with an initial state in which all fields are zero and there is a source, these can be then sequentially updated in a time stepping manner. The propagation of physical waves is then simulated through the advancement of the numerical fields.

Then, at the end of the simulation, the power density P_D across the computational domain can be obtained as the module of the Poynting vector S_P, which is obtained from the final values of the electric and magnetic fields as

$$S_P = E \times H. \tag{2.12}$$

Then, the power P captured by a given antenna is $P = A_e \cdot P_D$ where A_e is its effective area. Receiving antennas with different effective areas can thus be modelled by simply considering their gain and adding it to the power result.

2.1.0.1 Numerical dispersion

Although *dispersion* is commonly used as synonym of spread, in physics *dispersion* has a more restrictive definition. In the case of concern in this thesis, dispersion makes reference to the fact that the properties of a wave (e.g. phase velocity, wavelength, etc) depend on its frequency.

The relationships between these parameters and the frequency of an electromagnetic wave are well known by physicists and vary between different materials. The most representative of these relationships is the *dispersion relation*, which relates the wave vector to the frequency of the wave. For instance, the analytical dispersion

relation of a physical plane wave propagating within a homogeneous lossless material is

$$(k_x)^2 + (k_y)^2 + (k_z)^2 = \left(\frac{\omega}{c}\right)^2 \tag{2.13}$$

where ω is the angular frequency, c the speed of light in that medium and k_ζ with $\zeta \in \{x, y, z\}$ are the Cartesian components of the wave vector.

In addition, numerical waves propagating in a Yee grid according to the FDTD equations presented earlier, are also subject to dispersion. Further, the numerical dispersion relationship [31] in FDTD is

$$\sum_{\forall\zeta} \frac{1}{(\Delta\zeta)^2} sin^2\left(\frac{\tilde{k}_\zeta \Delta\zeta}{2}\right) = \frac{1}{(c\Delta t)^2} sin^2\left(\frac{\omega\Delta t}{2}\right) \tag{2.14}$$

with $\Delta\zeta$ the cell size along the ζ dimension, and $\zeta \in \{x, y, z\}$ in 3D or $\zeta \in \{x, y\}$ in 2D. This relation is clearly different from the physical one of equation (2.13) and hence, waves in numerical grids do not propagate in the same way that physical waves do.

The dispersive behaviour of the FDTD grid for narrowband simulations is also well known [31]. In particular, Zhao [32] characterised in 2003 the extrema of the numerical phase velocity, thus quantifying the anisotropy error in the propagation of numerical CW signals. This error can be reduced by making the spatial step small compared to the physical wavelength (in exchange for larger computational requirements). However, when simulations are performed at low spatial resolutions, the impact of dispersion grows and several numerical artifacts appear. The results of such a phenomenon are particularly disruptive in wideband simulations, which are the subject of study in chapter 6.

Finally, let us also define further grid parameters with influence in the dispersive behaviour of the grid. First, the cell aspect ratios are defined as

$$R_y = \frac{\Delta x}{\Delta y} \tag{2.15}$$

$$R_z = \frac{\Delta x}{\Delta z}. \tag{2.16}$$

Also, let the spatial grid resolution be defined with respect to the wavelength corresponding to the carrier frequency:

$$N_{\lambda_0} = \frac{\lambda_0}{\Delta x}. \tag{2.17}$$

And finally, let the normalised Courant stability factor (aka Courant number) be

$$S = \sqrt{\sum_{\forall \zeta} \left(\frac{c\Delta t}{\Delta \zeta}\right)^2} \tag{2.18}$$

with $\zeta \in \{x, y, z\}$ in 3D or $\zeta \in \{x, y\}$ in 2D. Further, let $S \in [0, 1]$ to guarantee the numerical stability of the algorithm, which is known as the Courant stability condition [31].

2.2 The convolutional perfectly matched layer

Residential scenarios and urban environments are "open" regions, i.e. they are not enclosed within a solid structure. However, their numerical representation in the FDTD computation is limited, thus inherently having a border. This boundary around the scenario acts as a perfect electrical conductor (PEC), hence reflecting outgoing energy back into the computational domain. Such an effect is undesired, as it introduces an artificial source of reflections in the simulation. Hence, an absorbing boundary condition (ABC) is necessary around the computational domain to avoid simulation errors due to reflections on the border.

In 1994, a major breakthrough in this problem took place when Berenger introduced the perfectly matched layer (PML) [33], an efficient numerical absorbing material with an impedance that is matched to waves of whatever polarisation, frequency and angle of incidence. This way, a special material can be placed around the computational domain (see Figure 2.2) such that it attenuates the energy leaving the scenario. Then in 2000, Roden and Gedney presented an efficient implementation of Berenger's PML called the CPML [34], which has since been the choice for this purpose.

The original PML formulation is rather tedious. Thus, in [35] a more elegant formulation was introduced that consisted on mapping Maxwell's equations into a complex stretched-coordinate space by making use of a complex frequency shifted (CFS)

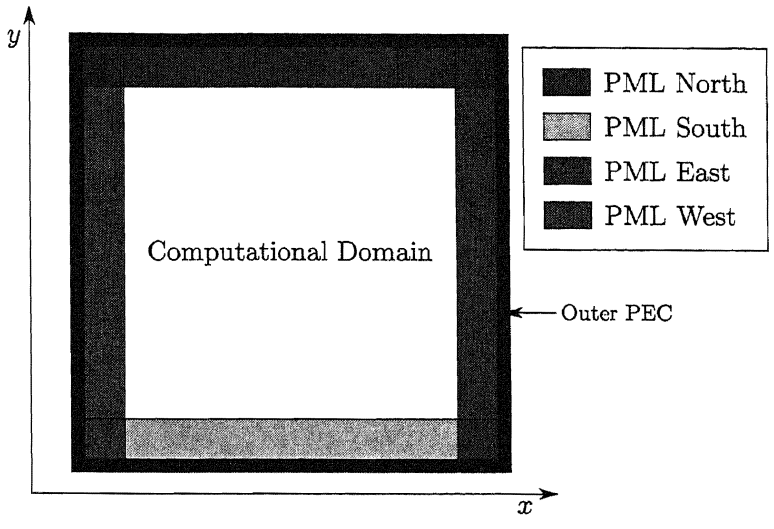


Figure 2.2: PML regions.

tensor. The formulation of this tensor was further generalised by Roden and Gedney in [34], thus leading to the one currently used in their CPML implementation:

$$s_w = \kappa_w + \frac{\sigma_w}{a_w + j\omega\varepsilon_0} \qquad w \in \{x, y\} \tag{2.19}$$

with $\kappa_w \geq 1$ and $a_w > 0$. Also note that σ_w is equivalent to the electric conductivity of the PML and is thus responsible for the losses within it. Then from [31] and assuming that the s_w tensors are continuous functions along the axial directions, the mapping from the standard coordinates x, y and z into the CFS coordinates $\tilde{x}$, $\tilde{y}$ and $\tilde{z}$ is

$$\tilde{x} \longrightarrow \int_0^x s_x(x')dx' \qquad \tilde{y} \longrightarrow \int_0^y s_y(y')dy' \qquad \tilde{z} \longrightarrow \int_0^z s_z(z')dz'. \tag{2.20}$$

The main idea beneath this approach is that *any coordinate transformation* of Maxwell's equations is equivalent to expressing them in traditional Cartesian coordinates with *transformed materials* that do not necessarily exist in the physical world. However, such materials are still useful in computer simulations for the absorption of outgoing energy.

In order to avoid reflections between the computational domain and the CPML boundary due to the discontinuity of s_w, the losses of the CPML must be zero at the interface with the computational domain. These losses are then gradually increased

in an orthogonal direction from the interface towards the outer PEC boundary. A polynomial grading of a_w, κ_w and σ_w has been shown [31] to be optimal for this task:

$$a_w(w) = a_{w,max} \cdot \left(\frac{d-w}{d}\right)^{m_a} \tag{2.21}$$

$$\sigma_w(w) = \left(\frac{w}{d}\right)^{m} \cdot \sigma_{w,max} \tag{2.22}$$

$$\kappa_w(w) = 1 + (\kappa_{w,max} - 1) \cdot \left(\frac{w}{d}\right)^{m} \tag{2.23}$$

where d is the depth of the CPML, m_a and m are the scaling orders and $0 \leq w \leq d$. In order to guarantee high attenuation of low frequencies by the CPML, it has been shown [31] that the largest a_w value must be $a_{w,max} << 2\pi f_a \varepsilon_0$ with f_a the lowest significant frequency of the transmitted signal.

Further, when polynomial grading of σ_w is used, the optimal $\sigma_{w,max}$ can be determined for a given reflection error $R(0)$ as

$$\sigma_{w,opt} = -\frac{(m+1)ln[R(0)]}{2\eta d} \tag{2.24}$$

where η is the impedance of the background material. Regarding, $\kappa_{w,max}$, it has been seen [31] that values between 7 and 20 reach reflection errors below -70 dB, which is more than sufficient for the applications under consideration.

Finally and following the notation of [31], the FDTD equations (2.6) to (2.8) can be modified for the implementation of the CPML, thus leading to

$$H_x\big|^{n+1}_{i-0.5,j+1} = H_x\big|^{n}_{i-0.5,j+1} - D_b\big|_{i-0.5,j+1} \cdot \left[\frac{E_z\big|^{n+0.5}_{i-0.5,j+1.5} - E_z\big|^{n+0.5}_{i-0.5,j-0.5}}{\Delta y \cdot \kappa_{y_{j+1}}} + \Psi_{H_x,y}\big|^{n+0.5}_{i-0.5,j+1}\right] \tag{2.25}$$

$$H_y\big|^{n+1}_{i,j+0.5} = H_y\big|^{n}_{i,j+0.5} + D_b\big|_{i,j+0.5} \cdot \left[\frac{E_z\big|^{n+0.5}_{i+0.5,j+0.5} - E_z\big|^{n+0.5}_{i-0.5,j+0.5}}{\Delta x \cdot \kappa_{x_i}} + \Psi_{H_y,x}\big|^{n+0.5}_{i,j+0.5}\right] \tag{2.26}$$

$$E_z\big|_{i-0.5,j+0.5}^{n+0.5} = C_a\big|_{i-0.5,j+0.5} \cdot E_z\big|_{i-0.5,j+0.5}^{n-0.5} + C_b\big|_{i-0.5,j+0.5} \cdot \Big[\Psi_{E_{z,x}}\big|_{i-0.5,j+0.5}^{n} - \Psi_{E_{z,y}}\big|_{i-0.5,j+0.5}^{n} + \frac{H_y\big|_{i,j+0.5}^{n} - H_y\big|_{i-1,j+0.5}^{n}}{\Delta x \cdot \kappa_{x_{i-0.5}}} - \frac{H_x\big|_{i-0.5,j+1}^{n} - H_x\big|_{i-0.5,j}^{n}}{\Delta y \cdot \kappa_{y_{j+0.5}}}\Big]. \tag{2.27}$$

Here, $\Psi_{H_{x,y}}$, $\Psi_{H_{y,x}}$, $\Psi_{E_{z,x}}$ and $\Psi_{E_{z,y}}$ are discrete variables with nonzero values only in their corresponding PML regions. According to the scenario partition of Figure 2.2, such regions are:

- $\Psi_{H_{x,y}}$ and $\Psi_{E_{z,y}}$: Nonzero only in the north and south regions of the PML.
- $\Psi_{H_{y,x}}$ and $\Psi_{E_{z,x}}$: Nonzero only in the east and west regions of the PML.

And the update equations of the auxiliary variables are:

$$\Psi_{H_{x,y}}\big|_{i-0.5,j+1}^{n+0.5} = b_{y_{j+1}} \cdot \Psi_{H_{x,y}}\big|_{i-0.5,j+1}^{n-0.5} + c_{y_{j+0.5}} \cdot \left(\frac{E_z\big|_{i-0.5,j+1.5}^{n+0.5} - E_z\big|_{i-0.5,j+0.5}^{n+0.5}}{\Delta y}\right) \tag{2.28}$$

$$\Psi_{H_{y,x}}\big|_{i,j+0.5}^{n+0.5} = b_{x_i} \cdot \Psi_{H_{y,x}}\big|_{i,j+0.5}^{n-0.5} + c_{x_i} \cdot \left(\frac{E_z\big|_{i+0.5,j+0.5}^{n+0.5} - E_z\big|_{i-0.5,j+0.5}^{n+0.5}}{\Delta x}\right) \tag{2.29}$$

$$\Psi_{E_{z,x}}\big|_{i-0.5,j+0.5}^{n} = b_{x_{i-0.5}} \cdot \Psi_{E_{z,x}}\big|_{i-0.5,j+0.5}^{n-1} + c_{x_{i-0.5}} \cdot \left(\frac{H_y\big|_{i+1,j+0.5}^{n} - H_y\big|_{i-1,j+0.5}^{n}}{\Delta x}\right) \tag{2.30}$$

$$\Psi_{E_{z,y}}\big|_{i-0.5,j+0.5}^{n} = b_{y_{j+0.5}} \cdot \Psi_{E_{z,y}}\big|_{i-0.5,j+0.5}^{n-1} + c_{y_{j+0.5}} \cdot \left(\frac{H_x\big|_{i-0.5,j+1}^{n} - H_x\big|_{i-0.5,j}^{n}}{\Delta y}\right) \tag{2.31}$$

where

$$b_w = e^{-\left(\frac{\sigma_w}{\kappa_w} + a_w\right) \cdot \frac{\Delta t}{\varepsilon_0}} \tag{2.32}$$

$$c_w = \frac{\sigma_w}{\sigma_w \kappa_w + \kappa_w^2 a_w} \cdot (b_w - 1). \tag{2.33}$$

These are the final equations used in section 3.3 to describe the implementation of the FDTD algorithm with CPML on a parallel computing platform.

2.3 General purpose computing on graphics processing units

In FDTD, the information within each grid cell is updated every time step. In addition, this algorithm does not require different computational flows for the update of field values at different locations. Hence, each parallel thread can be designed for the calculation of one single field cell. This way, all threads follow the same computational flow, which can be simultaneously applied over different cells. This makes of the FDTD algorithm a suitable candidate for implementation on parallel computing architectures [36]. For this reason, distributed computing approaches such as volunteer computing were also evaluated in the beginning of this research work. However, the spatial differentiation process of a given FDTD cell requires the use of information contained in the adjacent cells. Hence, distributed implementations of this algorithm have communication requirements that are proportional to the matrix size. Further, the homologous calculation process across cells removes the need for asynchronous thread execution provided by distributed approaches. Due to this, a non-distributed parallel computing model also satisfies the requirements for FDTD implementation.

Because of the reasons expressed above, two parallel hardware acceleration techniques have received most of the attention in recent years from the FDTD community. These are field-programmable gate-arrays (FPGAs) and graphics processing units (GPUs). For instance, [37] claimed in 2004 speeds of 75 Mcps[1] for a 2D implementation on an FPGA. However due to market demands, FPGAs tend to be more costly devices than GPUs, which exist nowadays in every personal computer, thus being more readily accessible for testing. In addition, graphics chipsets are becoming cheap and powerful, and their architecture is also well adapted to parallel algorithms. Hence, with the development of programmable graphics hardware, further solutions to compute electromagnetics on GPUs have arisen in the last years. For example, [38] makes use of a multi-resolution FDTD implementation on a GPU as a tool to model a microwave wireless channel. Another example of computational velocity is that of [39], which in 2009 achieved speeds of 450 Mcps on a *Quadro*

[1] Mega cells per second.

FX 5600 graphics card for a 3D implementation. Since GPU technology advances quickly, one attractive feature of this platform for scientific computing is that higher speeds can be easily achieved by simply replacing old platforms without redesigning the software. This is interesting for most companies, since it reduces their needs for continuous development.

Finally, the main advantages of GPUs for scientific computing can be summarised as follows:

- **GPUs are usually cheaper** than other platforms (e.g. computer clusters), for the same computational power.
- **GPUs are more accessible** than other devices such as for instance, FPGAs. In fact, there is usually one GPU in almost every PC.
- **The learning curve is usually softer** for RF engineers. This usually translates in a shorter development cycle, thus also saving costs.

For all these reasons, the work presented in this thesis focused on a GPU implementation of the FDTD algorithm, which led in 2008 to the publication one of the first FDTD implementations in a GPU using CUDA.

2.3.1 The Compute Unified Device Architecture

In 2007, the Compute Unified Device Architecture (CUDA) [40] was first released, thus providing a *C*-based interface to the programming of GPUs. Due to this, the learning curve for *C* programmers is softer than with specialised hardware such as FPGAs, which require specific architecture knowledge. Furthermore, the CUDA driver deals directly with the data transfer between GPU and CPU, thus releasing the programmer from having to implement low-level data transfer routines, which are prone to errors. Moreover, if the memory accesses can be coalesced, as it is the case with FDTD, memory reads with CUDA are capable of recovering several words in only one memory access, thus substantially increasing the speed of data transfer. Hence for all these reasons, a decision was made in the beginning of this research project to use GPU computing with CUDA for the implementation of the standard FDTD algorithm.

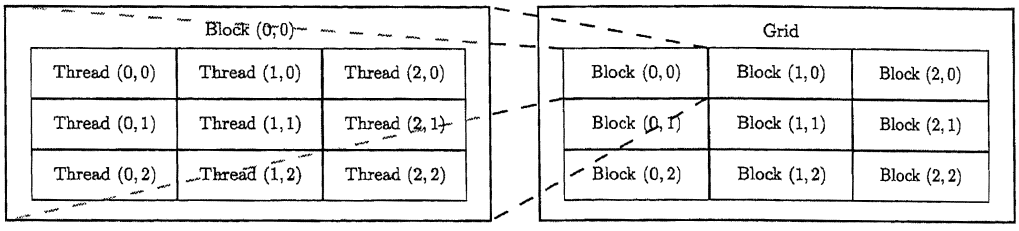


Figure 2.3: Threads hierarchy in CUDA.

Table 2.1: CUDA terminology

Thread	Sequence of executable instructions
Block	Arrangement of threads
Grid	Arrangement of blocks
Warp	Group of 32 consecutive threads from the same block
Half-warp	Either the first or the second half of a warp
Kernel	Programs that are parallelly executed

Similarly to other parallel programming platforms, the GPU is a device capable of executing several *threads* in a multithreading fashion. Further, a GPU contains several multiprocessors, each being responsible for executing threads parallely. This is done by the eight scalar processor cores contained within each multiprocessor. The memory model offered by CUDA is shared and the developer is responsible for writting *kernels*, which are the pieces of *C* code that each thread executes. The compiler then translates the kernels into the instruction set of the device for execution. Different threads are always bundled together into a thread *block* in a way that allows them to cooperate by exchanging data through shared memory. This is illustrated on Figure 2.3. Since one block contains only a limited number of threads, blocks can also be grouped together into a block *grid*, thus increasing the total number of threads that can be executed.

The multiprocessors within the GPU are then responsible for executing the different thread blocks. After receiving a block, the multiprocessor divides it into groups of 32 threads named *warps*. Then, warps are executed one common instruction at a time. Thus if some threads within a warp diverge, i.e. if they have to execute different instructions, their execution must then be serialised and the execution time of the warp grows. Hence, the maximum efficiency is obtained when

all threads within the same warp follow the same execution path. This is achieved through careful programming that avoids divergence within the same warp when conditional statements like *if* are evaluated. The ratio between the number of active warps to the maximum amount of warps that can be active simultaneously is called the *multiprocessor occupancy* and it is the responsability of the programmer to maximise it in order to exploit the hardware resources. It must be noted that the maximum number of warps that can be executed simultaneously depends on the device being used. Standard off-the-shelf GPUs have the necessary number of multiprocessors to support the requirements of most common users. However, professional graphic cards for scientific computing have more powerful GPUs with more and faster multiprocessors.

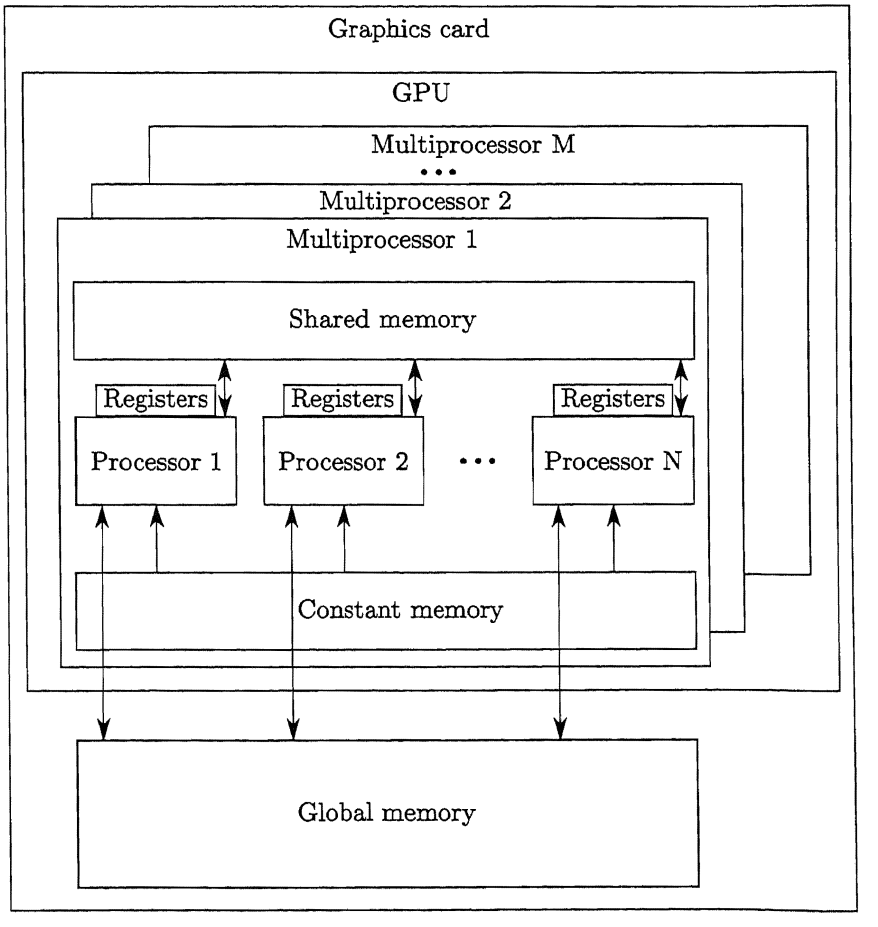


Figure 2.4: Device architecture.

This computational scheme is akin to the one found in single instruction - multiple data (SIMD) architectures. However, in CUDA terminology, this scheme is called single instruction - multiple thread (SIMT). According to [40], the difference lies in that, in SIMD arrangements, the programmer must choose the SIMD width for a group of parallel threads, regardless of wether the computing platform can parallelise them. However in SIMT, the width is fixed (32 threads per warp), thus allowing the programmer to write thread-level code and achieving higher parallelisation based on the knowledge of the warp size.

As seen, SIMD and SIMT schemes execute instructions parallely over several data words. Thus, the handling of large portions of data also represents an important burden to the computation. Nevertheless with CUDA, there are different types of memory spaces with different access patterns. It is thus important to design the memory access carefully in order not to slow down the transfer of data. The memory schemes and the general hardware structure are explained in the following and are also illustrated on Figure 2.4.

Each thread has its own memory space, which is termed as *local memory* and is only accessible by the tread itself for some automatic variables. Further, each block can also make use of *registers* for the storage of the variables defined in the kernels. The access to registers does not add extra clock cycles per instruction. However, some accesses may have to be put on hold in order to wait for previous reads or writes. In addition, there is a limited number of registers per multiprocessor. Hence the programmer must be careful not to use an excessive number of auxiliary variables per kernel. Otherwise, the number of blocks that can be executed simultaneously will be reduced.

Global memory represents the largest memory space on the graphics card and so it commonly deals with the largest portions of data transfer to/from the GPU. In the following, global memory is considered to be divided into logical *segments* of length 32, 64 and 128 bytes. Usually, the GPU reads 4, 8 and 16 byte words from global memory in only one instruction. However, the most efficient access method is by means of *coalesced* access. This is only possible for certain memory accesses and it means that the retrieval/storage of several byte words is grouped into a single and contiguous memory access. Coalesced reads/writes from/to global memory are possible in CUDA when the memory accesses by the threads in a *half-warp* fulfil

certain requirements. As long as these requisites are met and even if the threads within a warp diverge, memory accesses can be coalesced. These requirements vary slightly across computing platforms an are explained in depth in [40]. However, for the most common devices, these can be summarised as:

- Each thread must access words of 4, 8 or 16 bytes in length.
- All words must reside in the same memory segment.
- The access to each word by the threads must be sequential, i.e. the n^{th} thread accesses the n^{th} word.

Another memory space is *constant memory*, which is backed by an on-chip cache. Thanks to this, a read from constant memory accesses the global memory only on a cache miss. For all threads within a half-warp, this type of access is as fast as accessing a register.

Shared memory, which also resides on-chip, is another fast memory space. It is called shared because all the threads within the same block can use it to exchange information throughout synchronised reads and writes. The registers that make up the shared memory space are grouped in modules called banks and the access to different banks is done in parallel. However, if two different threads try to access the same bank, the accesses have to be serialised. It is thus important to design the shared memory accesses so as to avoid bank conflicts. Finally, there is also a *texture* memory space. However, CUDA does not currently support texture rendering (writing to texture memory) and hence, this memory space has not been used in the current work.

2.4 Measurements equipment

During the development of this thesis, a series of coverage measurements have been performed. These are necessary for the calibration and evaluation of the coverage prediction methodology to be introduced in chapter 4. Thus in this section, the different pieces of equipment that have been used for the obtainment of the empirical measurements are described.

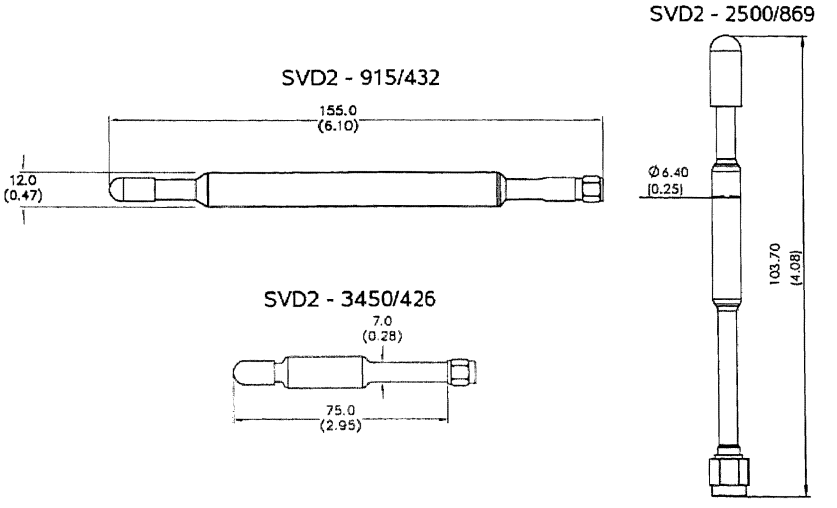


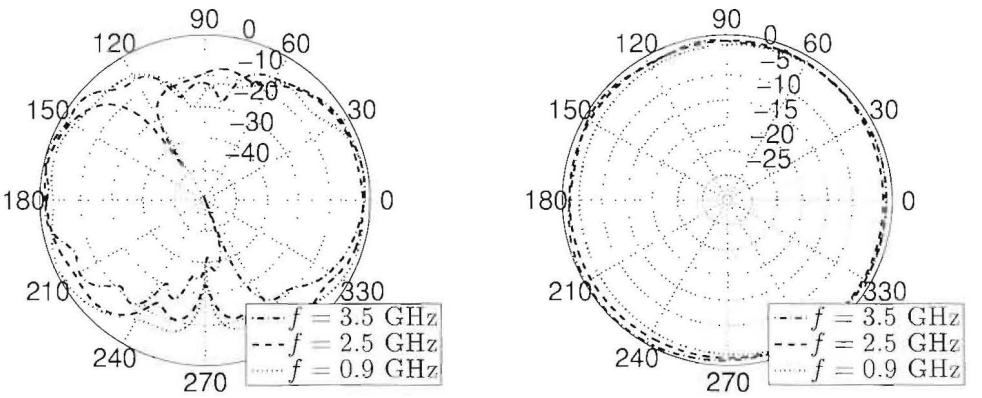
Figure 2.5: Dimensions of antennas with units in mm (inches). Diagrams extracted from *Cobham* data sheets.

Table 2.2: Antenna characteristics

Frequency range	Model	Max. Gain	Vert. HPBW	Max. VSWR
870 - 960 MHz	SVD2 - 915/432	2 dBi	80°	2 : 1
2.38 - 2.61 GHz	SVD2 - 2500/869	2 dBi	80°	2 : 1
3.40 - 3.65 GHz	SVD2 - 3450/426	2 dBi	80°	2 : 1

2.4.1 Antennas

The employed antennas were dipoles similar to those used in WiFi access points and FAPs. More specifically, they were slim flexible dipoles of the SVD2 series from *Cobham* [41] (previously *European Antennas*). They have semi-rigid covers and their structures and dimensions are illustrated on Figure 2.5. They all have omniazimuthal radiation patterns with a wide vertical beamwidth, i.e. the half-power beamwidth (HPBW) is 80°. The radiation patterns are also illustrated on Figure 2.6. In addition, these antennas are all vertically polarised and have a maximum gain of 2 dBi in the horizontal direction. These and other characteristics as well as the antenna model names are summarised in Table 2.2.



(a) Elevation.

(b) Azimuth

Figure 2.6: Antenna patterns.

Figure 2.7: Vector signal generator MG3700A from Anritsu.

2.4.2 Vector signal generator

The radiated power emanates from a vector signal generator. The specific model is the MG3700A from Anritsu [42] (see Figure 2.7), which supports a frequency modulation range between 250 kHz and 6 GHz. Then, for the performance of the coverage measurements, unmodulated CW signals were generated at the carrier frequencies of interest.

2.4.3 Spectrum analyser

The measurements equipment used for the recording of the received power was a digital spectrum analyser. The sequence of processes applied by such a device to an input signal is as follows:

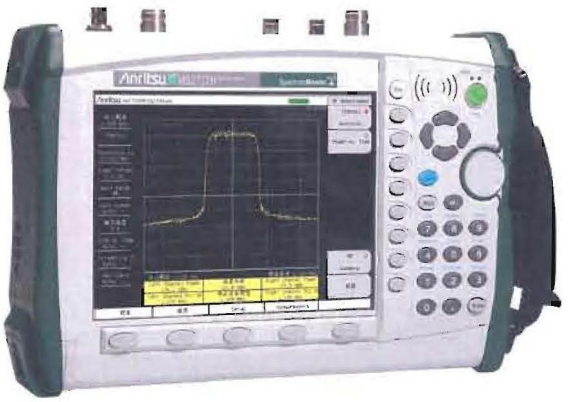


Figure 2.8: Spectrum analyser MS2721B from Anritsu.

1. The frequency of the input signal is first shifted to an intermediate frequency (IF) using a tunable oscillator.

2. Then, a tunable bandpass filter selects the band of interest.

3. The resulting signal is sampled using an analog-to-digital converter.

4. Finally, a discrete Fourier transform (DFT) of the resulting digital signal is computed.

The result of this process is thus an spectral power density which depends on the particular implementation of the DFT computation.

Different approaches for the final step of the spectrum analysis are necessary because the Fourier transform does not exist for all types of signals. For instance, not all wide sense stationary (WSS) processes have Fourier transforms. However, according to the Wiener-Khinchin theorem [43], the Fourier transform of a time-averaged autocorrelation function equals the power spectral density (PSD) for WSS processes. In fact, most non-vector spectrum analysers divide first a long signal into several segments, thus assuming time stationarity within each segment. These are next used as different realisations of the process. Autocorrelations can then be calculated and the power spectrum is obtained, although no phase information is displayed on screen.

The model used during the development of this work was the portable analyser MS2721B from Anritsu, which is illustrated on Figure 2.8. This device is small and

battery powered, which is convenient in order to perform handheld measurements while walking in order to average small-scale fading. Also, it contains a global positioning system (GPS) module, which can be used to record the location of each measurement.

2.5 Conclusions

This chapter has summarised the main tools used during the development of the research work presented in this thesis. These included the mathematical foundations of the finite-difference time-domain (FDTD) method for electromagnetic computation, as well as descriptions of its most important parameters. Then, the perfectly matched layer (PML) for FDTD simulations was introduced with the purpose of avoiding errors in the simulation of unbounded regions.

Further, a GPU-based parallel computing platform named the Compute Unified Device Architecture (CUDA) was also described. The concepts and terminology associated with this technology are important because they will be used to describe the implementation of a 2D FDTD simulator in the following chapter. Also, a set of hardware resources for the performance of empirical radio coverage measurements has been described. These included passive elements such as antennas, as well as a vector signal generator and an spectrum analyser.

All the concepts described in this chapter are intensely used in the following two chapters, where the scientific contributions of this dissertation are presented. To start with, chapter 4 studies first the increasingly common approach of low-frequency FDTD simulations, thus exposing its accuracy limitations for the purpose of coverage prediction. Then, a GPU implementation of the TM_z mode is described as a means of achieving fast execution times for the purpose of real-time wireless networks planning and optimisation (WNPO). Finally, the errors in the predictions of the described FDTD simulator will be assessed through a measurements-based calibration process.

3

Performance assessment

This chapter introduces a new approach to the coverage prediction of cellular base stations based on the finite-difference time-domain method for electromagnetic computation. Hence, an alternative to more traditionally used techniques such as ray tracing or empirical modelling is considered. As indicated in chapter 1, there have been previous attempts to use FDTD for such a purpose. However, the implementation techniques presented here allow for a substantially faster approach than previously reported.

For example, let us consider a problem related to the location and configuration of open access FAPs in an airport terminal with several candidate sites. To test different locations with various FAP setups, thousands of coverage predictions need to be performed. Hence, considering that network planning and optimisation is a dynamic process running along the lifetime of the network, coverage computation times measured on the order of hours or more are not acceptable. In addition, alternative existing ray-based solutions are already able of computing coverages in less than one minute [44]. As will be seen, the hardware-based implementation introduced in section 3.3 is capable of computing electrically large areas (i.e. residential femtocell scenarios) in seconds. This implementation is hence of interest, not only for the RF engineering community, but also for other areas of science seeking fast implementations of numerical FDTD solvers.

In the following, an overview of the computational complexity of the standard 2D FDTD is first presented. Then, an analysis of FDTD simulations at frequencies lower than the physical one is performed. This approach is used in the literature (e.g.

[45] or [46]) as a means to reduce the computation time. The reasons leading to such a technique are clarified in section 3.1 of this chapter. However, the implications of this method are not yet fully comprehended and they are thus studied in-depth in section 3.2 using a geometry-based analysis. Finally, an implementation of the algorithm that exploits its parallel properties and provides reasonable running times for radio channel modelling in large scenarios is introduced.

3.1 Complexity analysis

The complexity of the 2D FDTD is studied next in terms of the number of operations (time complexity) that this algorithm requires to solve an electromagnetic problem of a given size. Similarly, the memory requirements (spatial complexity) are also given. To do this, the number of field cells that must be computed to perform an electromagnetics simulation in a given area are calculated in the following. However, it is first necessary to define the input parameters of the algorithm.

Let the simulation scenario be of dimensions D_x and D_y. Further, let the simulated propagation duration be defined by the distance D_p that a phase of the wave is required to travel, i.e. a number N_t of time steps will be computed such that a phase of the wave covers the required distance D_p. The simulation frequency is f_0, the grid resolution along the x axis is N_{λ_0} and the normalised Courant number is S.

The cell sizes are then $\Delta x = c/(f_0 N_{\lambda_0})$ and $\Delta y = c/(R_y f_0 N_{\lambda_0})$, and the scenario size in number of cells is given by $N_x = D_x/\Delta x$ and $N_y = D_y/\Delta y$. Without loss of generality, $R_y \geq 1$ is assumed. Then, the total number of cells of *one* field component that are updated during *one* time step is

$$N_c = N_x \cdot N_y = D_x \cdot D_y \cdot R_y \cdot \left(\frac{f_0 N_{\lambda_0}}{c}\right)^2 \tag{3.1}$$

which, obviating the constants and using big-$\mathcal{O}$ notation, points out to the following spatial complexity:

$$\mathcal{O}(A \cdot R_y \cdot (f_0 N_{\lambda_0})^2) \tag{3.2}$$

with $A = D_x \cdot D_y$ the area of the scenario. Note however that, in a 2D simulation, there are three field components and hence, $3 \cdot N_c$ cells must be updated during each time iteration.

The time it takes for a phase of the wave to travel a distance D_p depends on its numerical phase velocity which, due to numerical dispersion, is different along different propagation directions. Considering the worst case, propagation along the direction with minimum numerical phase velocity $\tilde{v}_p$ is assumed. Then, a time period of duration $T = D_p/\tilde{v}_p$ is to be simulated. Further, the time step is given by

$$\Delta t = \frac{S\Delta x}{cR_s} \tag{3.3}$$

with $R_s = \sqrt{1 + R_y^2}$. Then, the total number N_t of time steps that need to be computed is

$$N_t = \frac{T}{\Delta t} = \frac{D_p R_s}{\tilde{v}_p S} f_0 N_{\lambda_0}. \tag{3.4}$$

Thus, the total number N_T of cells from *each* field component to be computed for the solution of this problem is given by the product of (3.1) and (3.4), which is

$$N_T = \frac{D_x D_y D_p R_y R_s}{c^2 \tilde{v}_p S} \cdot (f_0 N_{\lambda_0})^3. \tag{3.5}$$

It is seen from this equation that the computational load depends only linearly on the area of the scenario. Further, it is well known [31] that the numerical phase velocity tends to c, regardless of other parameters, as the spatial grid discretisation grows. Hence, $\tilde{v}_p$ can be merged with the c^2 factor in equation (3.5) to study the time complexity. In addition, the product $R_y R_s$ can be rewritten as $R_y \cdot \sqrt{1 + R_y^2}$ which, using big-$\mathcal{O}$ notation, has a growth of $\mathcal{O}(R_y^2)$. Finally and obviating the constants, this algorithm runs in

$$\mathcal{O}(A \cdot D_p \cdot R_y^2 \cdot (f_0 N_{\lambda_0})^3 / S) \tag{3.6}$$

time. This expression highlights that N_{λ_0} and f_0 are the parameters with more significant influence over the time complexity of the algorithm. The cells aspect ratio is secondary in weight over the complexity, altough in most FDTD implementations square grid cells are used ($R_y = 1$).

It is thus interesting, in order to save computational resources and lower the running time, to perform these simulations at the lowest frequency and spatial resolution possible. However, if N_{λ_0} is too low, the anisotropy error of the numerical phase velocity increases, thus undermining the simulation reliability. The consequences of FDTD simulations at low spatial resolutions are well understood [31]. Therefore, the following section investigates the accuracy of propagation modelling done with FDTD simulations at a frequency lower than the physical one.

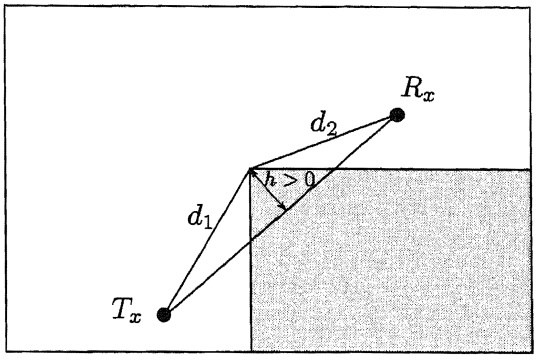


Figure 3.1: Plan view of a non-line-of-sight scenario.

3.2 Lower frequency simulations

A common approach to reducing the computational cost of finite-difference simulations is to perform them at a frequency lower than that of the original system. This procedure was applied, for instance, in [47]. As long as the spatial resolution remains constant, the spatial step in the simulation is then larger than at the original frequency. As a consequence of this, the matrix representing the computational domain contains less cells, thus occupying less memory. However, not only are memory savings achieved but also a cubic reduction in the number of necessary operations (see equation (3.6)). Then, a measurements-based calibration of the material parameters can be used to tune their electrical properties at the simulation frequency. Since the main propagation phenomena are approximated when the frequency is reduced, the calibration is imperative to reduce errors in the prediction. However, the geometry of the scenario still plays an important role on the radio propagation at different frequencies, in particular for diffractive effects. This consideration was never fully explored and its effect on the simulation results is thus studied in the following.

3.2.1 Theoretical attenuation due to knife-edge diffraction

In the most common scenarios of cellular communications, the receiver does not have a line-of-sight (LOS) to the transmitter. In such a case diffraction on the edges of buildings and at the indoor/outdoor interface is one of the main propagation phenomena.

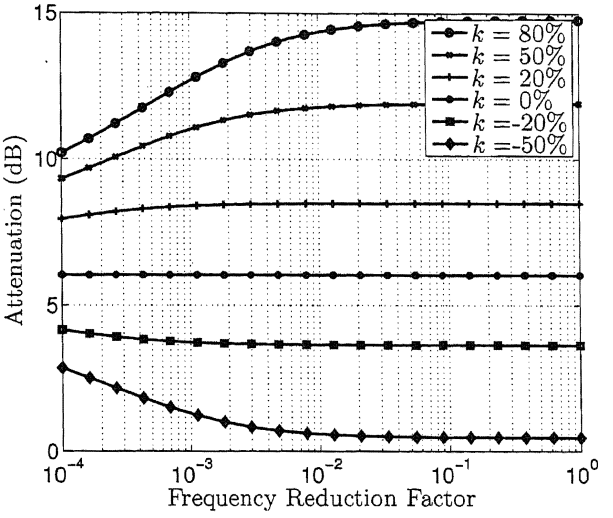


Figure 3.2: Theoretical attenuation due to diffraction when $f_0 = 3.5$ GHz. The obstacle is midway between transmitter and receiver, which are 20 meters apart.

When CW simulations are performed at a frequency lower than that of the physical carrier, the path loss varies. In particular, this is more notable for the losses due to energy absorption within materials as well as losses due to reflections. Nevertheless, these can be compensated throughout calibration of the properties of the different materials. However, diffractive effects are not so easily overcome due to the geometrical nature of diffraction itself. This means that the simulations performed at lower frequencies will be subject to an error because of diffractive effects. In order to estimate such an error, a prediction of the attenuation due to diffraction at lower frequencies is necessary. Recommendation ITU-R P.526-10 [48] provides already an empirical model for the evaluation of attenuation due to knife-edge diffraction and it is therefore taken as a reference for its calculation at different frequencies. In the following, the theoretical attenuation according to [48] will be characterised and compared to that measured from FDTD results at different simulation frequencies.

The simplest case is the study of the attenuation due to diffraction on the single edge of one obstacle (see Figure 3.1). In [48] h is defined as the size of the obstruction above the straight line linking transmitter and receiver. Since the formulas for the calculation of the attenuation are based on the Fresnel integrals, it is possible that

attenuation exists even if $h < 0$ due to the obstruction of the Fresnel zones.

To study the effect of using a lower frequency in finite-difference simulations on knife-edge diffraction over the edge of an obstacle, the frequency reduction factor (FRF) is defined as

$$\text{FRF} = \frac{f_{sim}}{f_0} \tag{3.7}$$

where f_{sim} is the frequency of the simulation and f_0 is the physical carrier frequency. Further, k is defined here as the ratio between the obstruction of the first Fresnel zone to its radius at the location of the obstacle:

$$k = \frac{h}{R_1}. \tag{3.8}$$

As mentioned earlier, the main purpose of using a lower frequency for FDTD simulations is to reduce the spatial step and the size of the matrix representing the computational domain. That is why usually $f_{sim} \leq f_0$ and this study is thus restricted to the values within the range $0 < \text{FRF} \leq 1$.

According to [48], the attenuation J due to diffraction on a knife edge at a given frequency is calculated with

$$J(\nu) = 6.9 + 20log\left(\sqrt{(\nu - 0.1)^2 + 1} + \nu - 0.1\right) \tag{3.9}$$

where

$$\nu = h\sqrt{\frac{2}{\lambda}\left(\frac{1}{d_1} + \frac{1}{d_2}\right)}, \tag{3.10}$$

λ is the wavelength and d_1 and d_2 are the distances to the edge as defined in Figure 3.1. Applying (3.10) and (3.9) to different values of FRF and calculating the attenuation for an obstacle with different k when FRF = 1, the attenuation values of Figure 3.2 are obtained. This highlights three different cases of diffractive effects:

1. If $k < 0$, there is a line-of-sight between transmitter and receiver but, depending on the simulation frequency, the obstacle can still obstruct the first Fresnel zone, thus introducing attenuation due to diffraction. To understand this phenomenon, let us consider the radius R_n in meters of the n^{th} Fresnel zone [48]:

$$R_n = 550\sqrt{\frac{nd_{T_x}d_{R_x}}{(d_{T_x} + d_{R_x})f}} \tag{3.11}$$

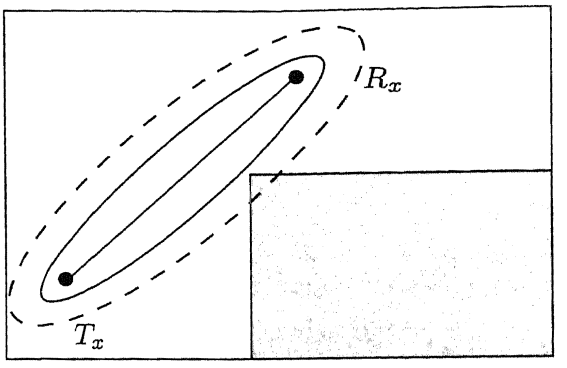


Figure 3.3: Line-Of-Sight propagation with diffractive effects.

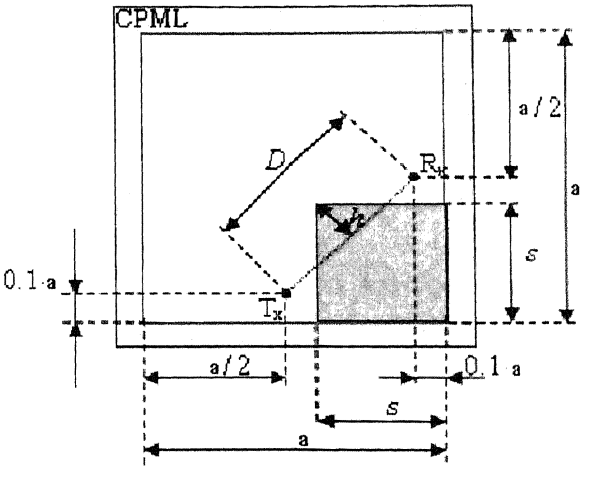


Figure 3.4: Test scenario for diffraction studies.

where f represents the frequency in MHz, and d_{T_x} and d_{R_x} are the distances in kilometers between transmitter/receiver and the point where the Fresnel radius is calculated. It can be seen from (3.11) that the use of a lower frequency increases the size of the Fresnel zone. The required volume for the radio propagation is thus bigger and the percentage of obstructed propagation volume increases with it. Figure 3.3 shows the change produced in the size of the Fresnel zone when decreasing the frequency (dashed line) and how different obstacles gain importance at lower frequencies, even in LOS conditions.

2. When $k = 0$, half of the first Fresnel zone is always obstructed, thus being the attenuation due to diffraction constant for any R_n and f_{sim}.

3. $k > 0$ is a non-LOS case in which half of the Fresnel zone is always obstructed. The remaining half of the ellipsoid will be more or less obstructed depending on R_n, i.e. the larger the radius, the lower the obstruction. The attenuation J due to knife-edge diffraction hence decreases at lower frequencies in agreement with Figure 3.2.

All of the above indicates that a reduction of the simulation frequency in FDTD can potentially lead to incorrect values of the field prediction due to unaccounted diffractive effects. Moreover, the error varies for different relative positions of transmitter and receiver, thus leading to anisotropic distortion:

- In line-of-sight cases with obstacles such as edges of houses, the predicted attenuation would be overestimated by the numerical simulation.
- In non-line-of-sight cases (more common) the attenuation due to diffraction would be underestimated by the numerical simulation.

3.2.2 Numerical attenuation due to knife-edge diffraction

To verify the previous theoretical results, the test scenario of Figure 3.4 has been designed. The dark gray square represents an obstacle with an edge that introduces diffractive effects. The size of this scenario is $a = 5$ m and the radius of the Fresnel zone midway between transmitter and receiver at $f_0 = 3.5$ GHz is $R_1 \approx 24.6$ cm.

Several FDTD simulations have been performed for values s of the side of the obstacle that result in the same degree k of obstruction as in Figure 3.2. The transmitter was a sinusoidal single point soft source, the grid resolution was $N_{\lambda_0} = 10$ and the normalised Courant number $S = 1$. This leads to a 584x584 matrix at $f_0 = 3.5$ GHz and smaller matrices at lower frequencies (FRF < 1). To simulate a reliable unbounded environment, the CPML introduced in section 2.2 has been used.

The attenuation due to knife-edge diffraction at the edge of the obstacle in the numerical simulation is then computed as follows:

1. An empty environment is first simulated and the attenuation A_{empty} at the position of the receiver is recorded. This attenuation is only due to free space propagation in the numerical grid.

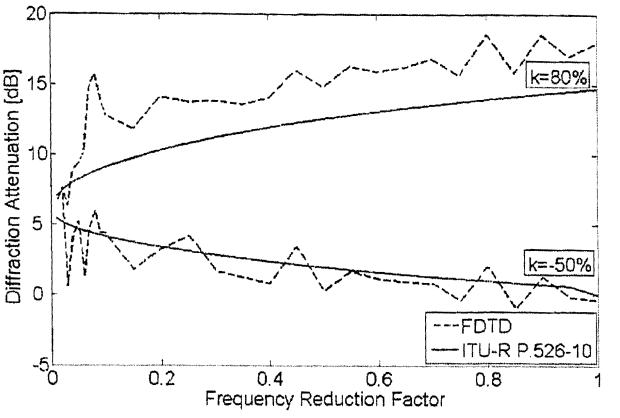


Figure 3.5: Numerical and theoretical diffraction predictions in the scenario shown in Figure 3.4 and described in section 3.2.2.

2. The simulation is repeated with the obstacle placed according to Figure 3.4 and the overall attenuation $A_{obstacle}$ is recorded.

3. The receiver location is chosen at a position where the main propagation phenomena are free-space and knife-edge diffraction. The attenuation in this case is then estimated with $J = A_{obstacle} - A_{empty}$.

In figure 3.5 the attenuation due to diffraction from the numerical simulations is displayed against the attenuation predicted by (3.9). The simulated obstructions in this case were $k = 80\%$ and $k = -50\%$ at the original frequency and it is observed that the numerical results show the same tendency as the theoretical ones.

It is also noticed that there exists a constant difference between the attenuation J measured from FDTD and the theoretical one for non-line-of-sight (NLOS) cases ($k = 80\%$). This is because the measured $A_{obstacle}$ does not account for the losses due to energy absorption. Since the size of the obstacle remains constant at all frequency reduction factors, the attenuation it introduces is approximately equal for all simulations. On the other hand, there is not such difference in the LOS cases, mainly because the LOS component is already strong enough. Despite of this, knife-edge diffraction is still the dominant propagation phenomena for the given scenario and the tendency of the numerical attenuation clearly agrees with the theoretical results.

3.2.3 Quantification of the error

An analytical evaluation of the error due to simulations with FRF < 1 can be deduced from the formulas given in [48]. Let us suppose for instance that simulation results are required for the physical carrier frequency $f_0 = 3.5$ GHz. Under the same settings used for the results of Figure 3.2, the attenuation due to diffraction on an isolated obstacle edge would be $J_0 \approx 14.72$ dB when the obstacle is obstructing $k = 80\%$ of the first Fresnel zone. If lower-frequency simulations with for example FRF $= 0.1$ were to be performed, the attenuation due to diffraction would be $J_{sim} \approx 9.11$ dB, thus giving rise to an absolute error of $E = J_0 - J_{sim} = 5.62$ dB. The physical attenuation due to diffraction is thus, in this case, underestimated by the numerical simulation. This error might not seem significant compared to other sources of attenuation such as free space propagation, reflection losses and/or building penetration. However, this may gain importance in more complex environments with several knife-edges.

By merging equation (3.7) with (3.10) and (3.9) a formula can be obtained for the evaluation of the error due to finite-difference simulations with FRF < 1:

$$\nu_0 = h\sqrt{\frac{2}{\lambda_0}\left(\frac{1}{d_1}+\frac{1}{d_2}\right)} \tag{3.12}$$

$$\nu_{sim} = \sqrt{\mathrm{FRF}} \cdot \nu_0 \tag{3.13}$$

$$E = 20log\left(\frac{\sqrt{(\nu_0 - 0.1)^2 + 1} + \nu_0 - 0.1}{\sqrt{(\nu_{sim} - 0.1)^2 + 1} + \nu_{sim} - 0.1}\right) \quad FRF \in (0, 1]. \tag{3.14}$$

A positive error indicates that the attenuation due to diffraction has been underestimated in the numerical simulation, while a negative error indicates an overestimation.

In Figure 3.6 the dependence of the error (3.14) with the frequency reduction factor and the obstruction k is presented. An interesting observation is that, in the worst case, i.e. when the FRF is the lowest possible (FRF $= 0$), the error in the prediction is limited and it depends only on the geometrical parameter ν_0 at the original frequency. This maximum value of the error could also be suspected from

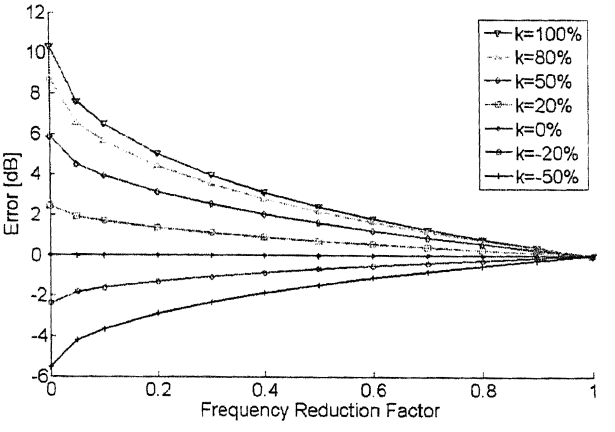


Figure 3.6: Attenuation error due to diffraction in the scenario shown in Figure 3.4 and described in section 3.2.2, when a lower frequency FDTD simulation is used. Note that the values for $FRF = 0$ are meaningless since there are no simulations at $f_{sim} = 0$.

Figure 3.2, where it is observed that the attenuation due to diffraction tends to a finite value as FRF decreases. Such value can be calculated by replacing $\nu = 0$ in (3.9) and its value is $J\,|_{\mathrm{FRF}=0} \approx 6.03dB$. This value arises from the formulation given in [48] and it is then straightforward to confirm from Figure 3.5 that the magnitude of the error will always be limited by $|E| \leq |J_0 - 6.03|$ with J_0 in dB.

In order to keep this error low, FRF must be as close to one as possible. However, the FRF is constrained by the size of the scenario to be simulated and the amount of memory available in the simulation platform. It is thus important to implement the algorithm in an efficient manner, such that its parallel nature is exploited for a fast execution time and the memory requirements are as low as possible. This allows the performance of simulations with high FRFs, thus keeping the error due to diffraction as low as possible. To achieve this purpose, the following section describes an efficient implementation of this algorithm.

3.3 Parallel implementation

As explained in the previous chapter and based on the existing literature, a parallel computing architecture is more appropriate for implementing the FDTD algorithm than using a traditional linear programming paradigm. Therefore this section describes an implementation of the 2D FDTD scheme on a GPU. The architecture

to be used is the Compute Unified Device Architecture (CUDA), which provides a shared memory parallel computing model and whose fundamentals were already introduced in section 2.3.1. In this scheme, memory management plays an important role in the efficiency with which data is loaded from and into the device's memory. In addition and to avoid instructions branching when evaluating conditional statements, separate kernels have to be designed for the evaluation of different regions within the computational domain. Hence, these and other implementation decisions are explained in the following.

The tool here presented is to be used in the rest of this thesis for the study of wireless channels and the FDTD model. Hence, the objectives to be fulfilled are

- To design and construct an FDTD simulator for the modelling of radio channels in femtocell scenarios.
- The running time of such simulator must be on the order of seconds for scenarios of the size of a femtocell coverage area and for a maximum physical carrier frequency of 3.5 GHz[1]. This is to allow network engineers to evaluate thousands of system-level simulations during the network planning and optimisation process.
- The frequency reduction factor must be as close to one as possible to avoid diffraction errors as explained in the previous section.

3.3.1 Kernels

One of the advantages of using a convolutional perfectly matched layer (CPML) for the simulation of unbounded scenarios is that it can be computed independently from the computational domain. This facilitates its implementation on an SIMT architecture because it eliminates the need for having to decide in each kernel, whether the current cell belongs to the absorbing layer or to the computational domain. Kernels are executed faster when all of the threads within a warp follow the same execution path. If the deciding condition evaluates differently for two or more threads within the same warp, the execution of the threads will have to be serialised, thus loosing the advantage of parallelisation. The parallel execution of threads that perform the

[1] 3.5 GHz is the frequency of WiMAX in Europe.

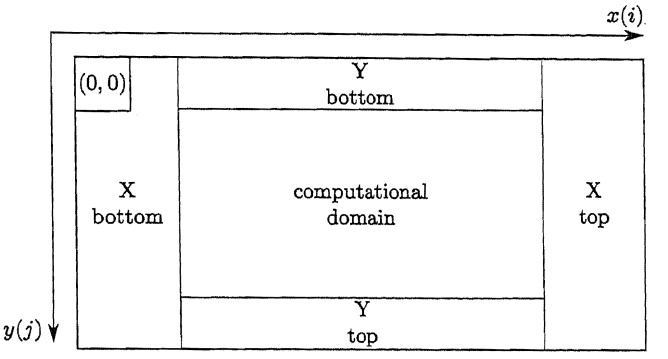


Figure 3.7: Division of the computational domain for kernel specialisation.

same task can be exploited by dividing the computation into different kernels for distinct parts of the scenario. According to this, five regions have been identified as illustrated on Figure 3.7. Furthermore, there are three types of fields (E_z, H_x and H_y) to be updated within each region, apart from the CPML auxiliary variables. Following then the notation of equations (2.25) to (2.31) for the field and CPML variables, the kernels that have been designed, compute the following variables:

1. $\Psi_{H_{x,y}}$ and H_x update in X-bottom (plus the corners for $\Psi_{H_{y,x}}$).
2. $\Psi_{H_{x,y}}$ and H_x update in X-top (plus the corners for $\Psi_{H_{y,x}}$).
3. $\Psi_{H_{y,x}}$ and H_y update in Y-bottom.
4. $\Psi_{H_{y,x}}$ and H_y update in Y-top.
5. H_x and H_y field update in the computational domain.
6. $\Psi_{E_{z,x}}$ and E_z update in X-bottom (plus the corners for $\Psi_{E_{z,y}}$).
7. $\Psi_{E_{z,x}}$ and E_z update in X-top (plus the corners for $\Psi_{E_{z,y}}$).
8. $\Psi_{E_{z,y}}$ and E_z update in Y bottom.
9. $\Psi_{E_{z,y}}$ and E_z update in Y top.
10. E_z field update in the computational domain.

The tasks that each kernel performs to update the FDTD fields, are:

1. Loading of the field values from the previous time iteration.

2. Computation of the new field values.

3. Storage of the result to memory.

In order to exploit the SIMT architecture, the workload is divided between all of the threads within each block. This is done by having each thread updating a single cell using the content of its neighbors. The source is updated on the host and copied to the corresponding grid cell in the card's global memory at each time step. It is thus not computed within the kernels. As an example, the pseudocode of the kernel that implements the electrical field update is shown in Algorithm 1.

Calculate the coordinates $i - 0.5$ and $j + 0.5$ of the current thread;
Load $H_x|^n_{i-0.5,j}$ into shared memory;
if *threadIdx.y* $= 0$ **then**
  Load $H_x|^n_{i-0.5,j+1}$ into shared memory;
end
Load $H_y|^n_{i,j+0.5}$ into shared memory;
if *threadIdx.x* $= 0$ **then**
  Load $H_y|^n_{i-1,j+0.5}$ into shared memory;
end
Synchronise threads;
Load material index;
Compute $E_z|^{n+0.5}_{i-0.5,j+0.5}$ according to (2.8);
Store result to global memory;

Algorithm 1: Electrical field update where *threadIdx* is a variable containing the thread indices (x and y) within the current block. Note that this is the code executed by one thread only.

3.3.2 Block and scenario size

The block size is also an important design parameter since larger blocks contain more threads. In this implementation, each thread loads one field value into shared memory (except for the E_z update kernel, in which H_x and H_y are loaded). There

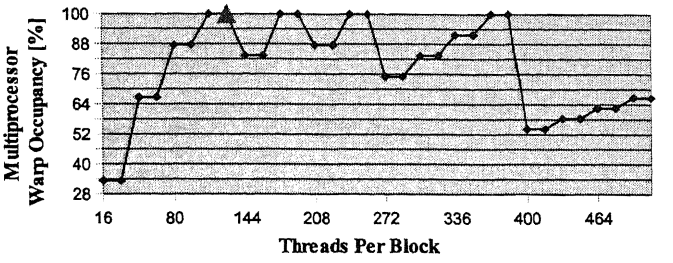


Figure 3.8: Multiprocessor occupancy

is thus a direct relationship between the block size and the amount of shared memory that a kernel needs. However, the amount of shared memory that each block can use is finite (see Table 3.1), so this imposes an upper limit to the block size. Furthermore, shared memory is shared between all the blocks that are executed on a multiprocessor. Considering these limitations, several numerical experiments indicated that a block size of $(B_x, B_y) = (16, 8)$ is optimal for this implementation in terms of data transfer speed. These are illustrated on Figure 3.8, which shows the effect on multiprocessor occupancy of various block sizes. The triangle indicates the multiprocessor occupancy reached with this block design, which corresponds to the highest multiprocessor occupancy. This configuration yields 128 threads per block and a multiprocessor occupancy of 100% for all kernels.

In order to maximise the instruction throughput by reducing divergence within warps, an efficient design is to match the block size to the geometry of the matrices. Ten cells are usually enough for the CPML to avoid waves reflecting back into the simulation scenario. However, since the maximum dimension of the block is 16, a CPML with thickness $N_{PML} = 16$ has been used, thus reducing instruction branching within warps and providing lower reflection coefficients. Therefore, the necessary number of cells along dimension $\zeta \in \{x, y\}$ for the discretisation of a 2D scenario of size (D_x, D_y) and with spatial steps $\Delta x = \Delta y = \Delta$ is

$$N_\zeta = \left\lceil \frac{D_\zeta}{\Delta} \cdot \frac{1}{B_\zeta} \right\rceil B_\zeta + 2N_{PML}. \tag{3.15}$$

Then, assuming that N_b bytes are necessary to store each cell, the memory requirements M_m in MB of one field matrix for such a simulation is

$$M_m = N_x \cdot N_y \cdot \frac{N_b}{1024^2} \quad (MB). \tag{3.16}$$

Table 3.1: Graphics cards used for computation

	GF 8600M GT	TESLA C870
Global Memory	256 MB	1.5 GB
Constant Memory	64 KB	64 KB
Shared Memory per Block	16 KB	16 KB
Clock Rate	337.5 MHz	1.35 GHz
Memory Bandwidth	9.6 GB/s	76.8 GB/s
Number of Multiprocessors	4	32

Considering that four field matrices must be read (E_z, H_x, H_y and the matrix containing the materials index) and three field matrices must be written (E_z, H_x and H_y) into global memory in each iteration, memory space for a total of seven matrices is necessary. Hence, the memory requirement M_f to store the field matrices is

$$M_f = 7 \cdot M_m = 7 \cdot N_x \cdot N_y \cdot \frac{N_b}{1024^2} \quad (MB). \tag{3.17}$$

In addition, memory is also needed for the auxiliary CPML variables in the PML regions. Following the previous approach, the amount of memory M_{CPML} in MB for these variables is:

$$M_{CPML} = 2 \cdot 4 \cdot (N_x + N_y) \cdot N_{PML} \cdot \frac{N_b}{1024^2} \quad (MB). \tag{3.18}$$

And finally, the total amount of memory required by this implementation is

$$M_T = M_f + M_{CPML} \quad (MB). \tag{3.19}$$

As mentioned above, each thread is responsible for computing one cell of the field matrices. The grid and block sizes must be therefore defined accordingly so that the total number of threads matches the total number of FDTD cells. In the cases where the size of the scenario is not a multiple of the block size, there will hence inevitably be some threads that do not need to compute anything, thus decreasing the total instructions throughput. That is why it is recommendable that the dimensions (N_x, N_y) in number of cells of the simulation scenario are as close as possible to a multiple of the block size. This can be easily done by adjusting the spatial step Δ. CUDA uses the Floating-Point Standard and the *float* type for

the representation of real numbers. Since the electrical and magnetic fields are real physical magnitudes, the value of a field corresponding to one cell can be stored as a float type in 4 bytes of memory. Therefore, this and the spatial step determine the memory requirements for the simulation of an scenario of a given size.

3.3.3 Memory access scheme

As seen above, the computation of each field consists mainly of some multiplications and additions, usually combined together by the compiler into a single fused multiply-add (FMAD) instruction. However, the other main tasks that kernels need to perform are the reading and writing of large matrices from/to *global memory*. This is the main bottleneck of the FDTD implementation so it is crucial to use an appropriate memory access scheme to obtain the maximum memory bandwidth.

Since it is not possible with CUDA to synchronise the execution of different thread blocks or to exchange information between them, each individual block has to load the update coefficients (material properties) from memory itself. Since there are usually 5 to 10 different materials in a simulation and $8 \cdot 16 = 128$ threads in a block, one block is enough to load this information. However, all blocks contain the same instructions sequence. Thus, having all blocks loading the material parameters is redundant. The impact of this redundancy in the memory access can be reduced by making use of the *constant memory* space. Furthermore and because this cache is physically located on each multiprocessor, almost all thread blocks will read from the cache once the parameters have been loaded (usually by the first block to be executed).

As seen in the previous chapter, another fast memory space is *shared memory*, which also resides on-chip. Since the update of each cell in an FDTD algorithm is based on the previous values of the fields in its neighbourhood, each cell needs to be accessed at least twice in each iteration. In order to minimise access to global memory, the threads of a block first load the corresponding submatrix into shared memory (see Algorithm 1), and later use these values to perform the computation. This way, the access to each cell in the global memory space is performed only once per time iteration.

As explained in section 2.3.1, some accesses to global memory can be *coalesced* into a single and contiguous memory access, as long as certain requirements are met. For instance, it was seen that each memory segment can have sizes of 8, 16 or 32 words of 4 bytes each and that coalesced access takes place when all threads access the same memory segment. Since in this design each thread is mapped to a matrix cell, the access to 2D matrices stored in global memory can be coalesced when the block width is a multiple of 8, 16 or 32. Further, from Figure 3.8, it is seen that the multiprocessor occupancy is maximum when the block size is $(B_x, B_y) = (16, 8)$ (128 threads), thus supporting the simulation of scenarios with widths that are a multiple of 16.

3.3.4 Algorithm performance

To illustrate the performance of the described implementation, a femtocell coverage prediction has been performed. For benchmarking reasons, two different graphics cards have been used to perform the computation: an off-the-shelf and non-expensive laptop graphics card (GF 8600M GT) and a high-performance computing card (TESLA C870). Their characteristics are summarised on Table 3.1.

The scenario is a residential area in a medium-size British town and with a size of approximately $(D_x, D_y) = (188, 199)$ meters. Using a spatial step of $\Delta = 12$ cm the scenario can be sampled into a matrix of size $(N_x, N_y) = (992, 1568)$ (roughly 1.5 million cells). However, the original scenario was slightly smaller. Since the width and height of a block of threads are 16 and 8 respectively, the environment was extended by adding some extra cells containing air in both dimensions. This way, the final matrix size is obtained. After adding the CPML cells, the total size of the matrix to simulate turns into $(N_x, N_y) = (1056, 1632)$. Thus, since each cell is stored as a floating point number in 4 bytes, each matrix occupies $4 \cdot N_x \cdot N_y \approx 6.7$ MB.

Since FDTD is a memory-intensive algorithm, memory bandwidth is an important feature to seek in a GPU for this purpose. As said earlier, seven large matrices must be read/written from/to global memory in each iteration. This amounts, in this simulation, to approximately 46 MB of data to be allocated in each time step.

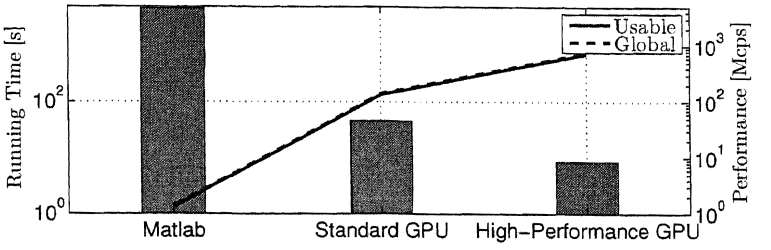


Figure 3.9: The bars represent the running time (left axis) of the algorithm with the setup described in section 3.3.4. The linear plots (right axis) represent the simulation speed in Megacells per second for the global scenario (including the PML) and for the computational domain only (usable part without the PML).

From equation (3.6) it is easily tested that this number grows quickly for larger scenarios or grid resolutions.

In order to measure the running time of this implementation for a matrix size of $(N_x, N_y) = (1056, 1632)$, a total of $N_t = 4000$ FDTD iterations have been performed. This is sufficient for the fields to propagate through the whole scenario after several reflections. However, the importance of simulating a large N_t lies on the increased statistical significance in the calculated computation time per cell. Using then the first GPU in Table 3.1, the total computation time has been of 46 seconds, which compares favourably with other existing propagation models. This algorithm has also been implemented in Matlab, which makes use of the AMD Core Math Library (ACML) for efficient matrix operations. The Matlab code was run on a computer equipped with an AMD Athlon 64 X2 Dual Core Processor 4600+ at 2.41 GHz and 3.25 GB of RAM. Applying the same configuration parameters as with the GPU simulation, the total computation time in Matlab was of approximately 78 minutes. This indicates that for the implementation of FDTD-based algorithms a standard off-the-shelf graphics card for laptops like the GeForce 8600M GT provides an speedup of around 100 times over highly optimised Linear Algebra libraries. Furthermore, when the computation is run on the high-performance computing card (TESLA C870), the simulation time is reduced to 8.6 seconds. With respect to the Matlab computation the speedup of TESLA is of 540. Figure 3.9 compares the measured running times as well as the performance of this FDTD implementation.

As seen, this implementation reaches speeds on the order of 800 Mcps on a

TESLA C870, which is sufficient for the purposes of channel modelling and WNPO initially described. However, during the last year this card has been superseeded by others with more memory, higher memory bandwidths and with more multiprocessors. It is thus expected that the running time of FDTD simulations with this implementation can be further reduced on more modern graphics cards.

3.4 Conclusions

This chapter has introduced the design considerations that must be taken into account when using CUDA for the implementation of finite-difference algorithms. An efficient memory access scheme has been introduced that matches the execution threads to the cells within the Yee lattice. Furthermore, the applicability of each memory space to the implementation of an FDTD algorithm to minimise memory accesses and increase simulation speed has been introduced. Besides and in order to minimise threads serialisation and exploit the parallel capabilities of the GPU, a subdivision of the computational domain and the absorbing boundaries to minimise instructions branching within threads of the same warp has been presented. Further, the tasks assigned to each execution thread have been explained and the dimensions of each thread block have been carefully shaped.

In addition to this, the reduction of the simulation frequency in finite-difference simulations has been studied through a novel geometric analysis based on Fresnel ellipsoids. It has been shown that this approach is subject to errors in the prediction of the attenuation due to diffraction at knife-edges. Further, it has been confirmed that knife-edge diffraction can model the effects suffered by the numerical wave when it encounters the edge of an obstacle, and the value of such attenuation has been evaluated. Additionally, it was shown that the attenuation values due to this phenomenon can be predicted using recommendation ITU-R P.526-10 on "Propagation by Diffraction" [48]. As a consequence of this, it can be said that the reduction of the simulation frequency is subject to a diffraction error due to the increase of the electrical size of the wave and of the radius of the Fresnel zone. This error can not be easily overcome through the calibration of materials because it depends on the geometry of the simulated scenario.

3.4.1 Contributions

The main contributions of this chapter made by the author to the scientific community are summarised in the following. In addition, numerical examples are also included for illustration purposes.

1. **Characterisation of the error due to lower-frequency simulations:** The analysis introduced here has highlighted that a measurements-based calibration is not enough to cope with diffraction errors due to finite-difference simulations at a lower frequency. Furthermore, such an error has been quantified and a formula has been devised to predict it.

   For instance, let us consider an FDTD coverage prediction at $f_0 = 3.5$ GHz. Due to limitations in the computational platform, a simulation frequency of $f_{sim} = 350$ MHz will be used ($FRF = f_{sim}/f_0 = 0.1$). As previously explained, lower-frequency simulations predict wrong attenuation values due to knife-edge diffraction. For example, consider an obstacle midway from transmitter and receiver that obstructs 50% of the first Fresnel zone at the physical carrier frequency f_0. Then, equation (3.14) showed that FDTD predicts an attenuation due to knife-edge diffraction 3.9 dB below the physical one. Consequently, diffraction predictions are unreliable.

2. **Efficient GPU-based implementation:** The parallel nature of the FDTD algorithm has been exploited for its implementation on a parallel hardware architecture. A novel execution flow and memory management scheme for such a design have been introduced. Besides and since the GPU market is large and heterogeneous, the guidelines provided here facilitate the selection of a proper GPU architecture for the solution of electrical problems of a given size.

   For instance, let us consider a scenario of size 100×50 meters to be simulated at $f_{sim} = 900$ MHz. Let us use a sampling step ten times smaller than the physical wavelength ($N_{\lambda_{sim}} = 10$) and a Courant number of $S = 0.9$. Furthermore, let us perform the computation in the TESLA C870 card described in Table 3.1. Since the implementation described in this chapter reaches a computational speed of about 800 Mcps, it is then seen that the described scenario is computed in approximately 9.4 seconds.

The analyses introduced in this chapter deal mainly with the reduction of the computation time of FDTD implementations. However, such FDTD predictions are not necessarily certain. To assess this question, the following chapter introduces a set of techniques that improves the accuracy of the predictions.

4

Coverage prediction with the finite-difference time-domain

FDTD algorithms are well-known for being computationally demanding. Today, they are commonly used in simulation scenarios of reduced size for applications such as the design of antennas [49] and microwave circuits as well as in studies of photonics [50]. However, even in such electrically small scenarios, the demands for a higher performance and a reduction of the computation time are regular. Thus, in order to apply this type of algorithms to the modelling of larger problems such as a street or an indoor environment for femtocells, efficient implementations are crucial. This was the main focus of the previous chapter, which introduced a novel parallel implementation of the standard FDTD.

Nevertheless, deterministic numerical simulations should be considered as mere approximations to reality. Furthermore, they are based on scenario descriptions that do not necessarily capture all the details of a physical site. Therefore, the reliability on these coverage predictions is only guaranteed through comparison with on-site measurements of radio propagation. Therefore, received power measurements have been performed. These have been used for the calibration and evaluation of the FDTD simulator described in chapter 3 and are described in section 4.1. Then, section 4.2 introduces a calibration methodology to configure the material properties such that the simulation result resembles the reality as faithfully as possible. Then, the model predictions are validated by comparison with additional measurements on a different site.

Table 4.1: Measurements settings

Parameter	Value
Carrier frequencies	0.9, 2.5, 3.5 GHz
Transmitter power output	19 dBm
Transmitter height	77 cm
Tilt	0°
Receiver height	$\approx$ 120 cm

4.1 Measurements of received power

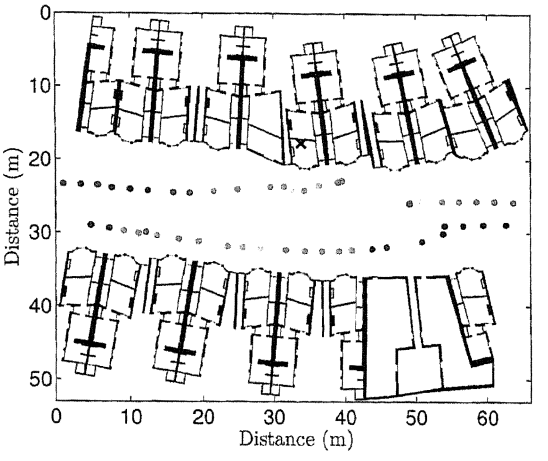
In order to evaluate the accuracy with which the previously described FDTD implementation can model radio propagation in the physical world, empirical measurements are necessary. These can then be used both for performance evaluation as well as inputs of the measurements-based calibration process to be described in section 4.2. Hence in the following, the set of measurements performed during the development of this thesis is presented. These consist of the recording of the received power in two femtocell-like scenarios and at three different carrier frequencies.

4.1.1 Equipment

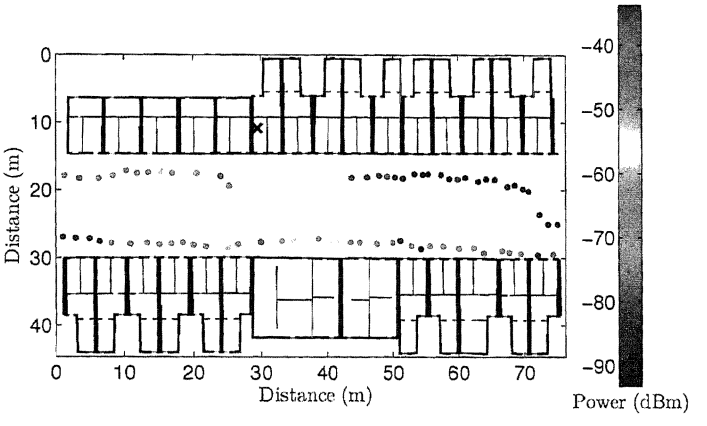
The transmitter, with an output power of P_{tx} = 19 dBm, is the vector signal generator MG3700A from Anritsu, which was described in section 2.4.2. Even though FAPs radiate between 10 and 15 dBm, this higher power level allows the recording of measurements at further locations from the transmitter. This way, the received signal is most of the times well above the noise floor, hence also increasing the reliability of the measurements at shorter distances. In order to perform CW power measurements, a continuously oscillating signal at the surveyed frequencies (900 MHz, 2.5 GHz and 3.5 GHz) was used. In addition, the antenna models used for the measurements were those described in section 2.4.1. Finally, the parameters of the measurements setup are summarised on Table 4.1.

4.1.2 Locations

The measurement sites were two residential neighbourhoods in a medium-size British town (see Figure 4.1). Femtocells leak power outdoors mainly through the premises



(a) Site 1.



(b) Site 2.

Figure 4.1: Indoors-to-outdoors power measurements at 2.5 GHz.

front facade (doors and windows), thus causing interference to nearby macrocells in the adjacent street-canyon of the residential area. Hence, the indoor-to-outdoor measurements have been performed in the street immediately adjacent to the premises and until distances from the outer wall at which the received power drops below -92 dBm (3 dB security margin from the noise floor at -95 dBm).

4.1.3 Measurements procedure

The measurements consisted, for each location, on the frequency sweep of a 1 kHz band centred at the previously described frequencies. This was done with a portable

spectrum analyser MS2721B from Anritsu, described also in section 2.4.3. The resolution and video bandwidths were 9 Hz, which is the minimum allowed by the device. Then, the received power P_{rx} is obtained as the largest power value within the scanned frequency band. The measurements device was carried by hand and the data was recorded while walking to reduce small-scale fading effects. Further, the locations of the measurements were resolved using a GPS receiver.

Finally, Figure 4.1 illustrates the surveyed areas along with some of the measurements.

4.1.4 Measurements postprocessing

Raw power measurements are subject to random fading phenomena and in particular to small-scale fading. Therefore in order to use the measurements for the calibration of this model, the data must first undergo a postprocessing phase during which outliers and fading effects are removed. Such postprocessing is detailed next:

Removal of location outliers: The location of the outdoor measurement points has been obtained using GPS coordinates, which are eventually subject to errors. Therefore, any measurement subject to the following location errors is to be removed: Out of range GPS coordinates, coordinates inside of a building, no GPS coverage or coordinates outside of the scenario.

Removal of noise bins: In areas of low coverage, it is possible that the measured signal becomes indistinguishable from the background noise. Those measurements are thus also classified as outliers. In order to clearly distinguish signal from noise, a large recording of the noise in the examined frequency band and location area has been performed. This way, the mean noise level has been found to be $\overline{N} = -95$ dBm. Then, any power measurement below -92 dBm (security margin of 3 dB from $\overline{N}$) is considered an outlier.

Small-scale fading removal: An averaging filter has been applied to the measurements for the removal of small-scale fading effects. This is done using an spatial filter that computes the average between all points within a distance of $40 \cdot \lambda_0$, i.e. the so-called 40-Lambda averaging criterion [51].

Table 4.2: Properties of materials commonly used in propagation predictions in the range (0.9 – 3.5) GHz.

Material	ε_r	μ_r	n	v_p (m/s)
Air @ STP	1.00058986 [54]	1	1.000295	299,704,079
Plaster	2.5 [55]	1	1.581139	189,605,399
Glass	6.38 [53]	1	2.525866	118,688,971
Wood	2.07 [53]	1	1.438750	208,370,162
Concrete	2.14 [53]	1	1.462874	204,933,905

After applying all the above described processes, the data can be then used for the purpose of calibration of propagation models.

4.2 Calibration of materials

In FDTD, the parameters characterising each material and thus playing an active role in the final simulation result are:

- The numerical relative electric permittivity $\tilde{\varepsilon}_r$.
- The numerical relative magnetic permeability $\tilde{\mu}_r$.
- The numerical electrical conductivity $\tilde{\sigma}$.

Just as before, the tilde ˜ identifies variables with significance only in the numerical domain. On the other hand, ε_r, μ_r and σ represent the properties of physical materials. For instance, Table 4.2 lists common values for some materials to be used in coverage predictions at radio frequencies. Since all of these are non-magnetic materials, $\mu_r = 1$ is generally assumed [52] [53].

However, due to the inherent uncertainties of the FDTD model under discussion (2D approximation, scenario description, etc), it should not be expected that the physical values of these parameters perform realistically in numerical simulations. This is particularly true when simulations are performed at a frequency lower than the physical one. These must thus be chosen carefully in order for the simulation result to resemble faithfully the reality. As advanced earlier, this can be achieved by using measurements of received power to find the combination of parameters that

Table 4.3: Physical constants

Constant	Value
ε_0	$8.85418781762 \cdot 10^{-12}$ F/m
μ_0	$4\pi \cdot 10^{-7}$ N/A^2
c	299,792,458 m/s

better match the prediction to the measurements. In the following, the process to decide on the values of $\tilde{\varepsilon}_r$ and $\tilde{\mu}_r$ is described.

The parameters ε_r and μ_r of a material, define its refractive index

$$n = \sqrt{\varepsilon_r \cdot \mu_r} \tag{4.1}$$

which is related to the phase velocity v_p of a monochromatic wave travelling through that material according to

$$v_p = \frac{c}{n} \tag{4.2}$$

where $c = 1/\sqrt{\varepsilon_0 \cdot \mu_0}$ is the speed of light in a vacuum (see Table 4.3 for the values of these physical constants). In the physical world (i.e. in nature) different values of ε_r and μ_r give rise to different phase velocities, being the same true for numerical simulations. In addition, the phase velocity of a monochromatic wave (both physical and numerical) can also be defined in terms of its angular frequency ω and wave number k as

$$v_p(\omega) = \frac{\omega}{k(\omega)}. \tag{4.3}$$

The FDTD scheme under study does not model dispersive materials, i.e. the refractive index is not frequency dependent. This may lead to the incorrect assumption that the numerical phase velocity $\tilde{v}_p$ is constant and fully defined by $\tilde{\varepsilon}_r$ and $\tilde{\mu}_r$ through the numerical equivalents of equations (4.1) and (4.2). However, as seen in section 2.1.0.1, the standard FDTD grid is dispersive and thus, $\tilde{v}_p$ varies for different propagation directions. This well understood phenomenon [31] gives rise to the physical phase-velocity error

$$\Delta\tilde{v}_{p_{physical}} = \frac{max\{|\tilde{v}_p - v_p|\}}{v_p}, \tag{4.4}$$

which according to [31], quantifies "the phase lead or lag that numerical modes suffer relative to physical modes propagating at v_p". Furthermore, since $\tilde{\varepsilon}_r$ and $\tilde{\mu}_r$ do not contribute extensively to the path loss, their values can be chosen prior to simulation so that the physical phase-velocity error is minimised.

For the purpose of coverage prediction, only monochromatic (single-frequency) simulations are to be performed. Then, for a given grid resolution, it is straightforward to adjust $\tilde{n}$ for each material such that (4.4) is minimised. This is done applying the strategy of [56] for centring a specific numerical phase velocity curve about a physical phase velocity v_p. To do that, the maximum, minimum and average values of $\tilde{v}_p$ with $\tilde{n} = 1$ are necessary. From [57], these are

$$(\tilde{v}_p)_M = \frac{\pi c}{N_\lambda R_s arcsin\left(\frac{1}{S} sin\left(\frac{\pi S}{N_\lambda R_s}\right)\right)} \tag{4.5}$$

$$(\tilde{v}_p)_m = \frac{\pi c}{N_\lambda arcsin\left(\frac{R_s}{S} sin\left(\frac{\pi S}{N_\lambda R_s}\right)\right)} \tag{4.6}$$

$$(\tilde{v}_p)_{avg} = \frac{(\tilde{v}_p)_m + (\tilde{v}_p)_M}{2} \tag{4.7}$$

where

$$R_s = \sqrt{1 + R_y^2 + R_z^2}. \tag{4.8}$$

For the FDTD scheme under study in 2D ($R_z = 0$) and with square grid cells ($R_y = 1$) it results that $R_s = \sqrt{2}$. Furthermore, the Courant number is chosen close to the stability limit ($S = 0.95$) to minimise the number of time steps. The numerical phase velocity $\tilde{v}_p$ is plotted on Figure 4.2, where it is seen that the average is far from being centred at v_p, thus being the physical phase-velocity error not optimum. Then, $\tilde{\varepsilon}_r$ and $\tilde{\mu}_r$ can be adjusted to minimise $\Delta\tilde{v}_{p_{physical}}$. Since the materials involved in these simulations are not magnetisable, $\tilde{\mu}_r = 1$ is chosen. Then, the $\tilde{\varepsilon}_r$ that produces an average $\tilde{v}_p$ equal to the physical phase velocity v_p of the material is

$$\tilde{\varepsilon}_r = \left(\frac{(\tilde{v}_p)_{avg}}{v_p}\right)^2 \tag{4.9}$$

Taking these criteria into account, the final values of the parameters ($\tilde{\varepsilon}_r$ and $\tilde{\mu}_r$) used in the coverage predictions of the following sections are those displayed on Table 4.4.

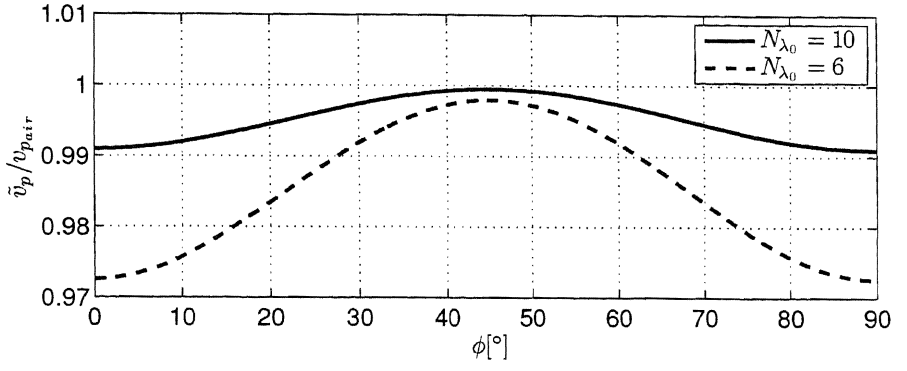


Figure 4.2: Normalised numerical phase velocity in 2D FDTD with square grid cells, $S = 0.95$ and $\tilde{n} = 1$. Normalisation is calculated with respect to the phase velocity through air: $v_{p_{air}} = c/n_{air}$ with n_{air} from Table 4.2.

Table 4.4: Material parameters for numerical coverage simulation with $S = 0.95$ to mimic the behaviour of the materials described in Table 4.2

Material	$\tilde{\varepsilon}_r$		$\tilde{\mu}_r$
	$N_{\lambda_0} = 6$	$N_{\lambda_0} = 10$	
Air	0.971	0.991	1
Plaster	2.425	2.475	1
Glass	6.190	6.315	1
Wood	2.008	2.049	1
Concrete	2.076	2.118	1

The refractive indices of the materials play a crucial role in wideband CIR simulations. This is because the time of arrival (ToA) of the received multipath pulses is determined by the numerical group velocity $\tilde{v}_g$, which is affected by $\tilde{n}$. Such an issue will be studied in depth in chapter 6. However, for the current case of coverage prediction, the values of $\tilde{\varepsilon}_r$ and $\tilde{\mu}_r$ are responsible only for the phase delays and thus for the resulting fading patterns. Meanwhile, the power distribution over the scenario (coverage prediction) is given by the scenario description itself and the attenuation introduced by each material, which is given by the numerical electrical conductivity $\tilde{\sigma}$. This is the remaining parameter to be tuned in order to guarantee the accuracy of the coverage prediction and is thus dealt with in the following.

4.2.1 Formulation of the optimisation problem

This section introduces a metaheuristics-based calibration approach that finds near-optimal values for the numerical electrical conductivity $\tilde{\sigma}$ of each material, such that the difference between power measurements and numerical predictions is minimised. These values can be later reused for further simulations in similar scenarios and under the same simulation conditions (grid resolution, time step, etc).

4.2.1.1 The materials error function

If $\tilde{\sigma}_m$ represents the electrical conductivity of material m, the s^{th} solution[1] to a problem involving N_m materials can be defined as

$$\Omega^s_{N_m} = \bigcup_{m=0}^{N_m-1} \tilde{\sigma}_{m,s}. \tag{4.10}$$

To evaluate the accuracy of a prediction using this solution, let us consider the measured power value P_{mes_p} in dBm at each measurement point p, with $p \in [0, N_p - 1]$ and N_p the number of points. Similarly, for an FDTD simulation using the conductivity values of $\Omega^s_{N_m}$ the same point is assigned a predicted power value $P^s_{pred_p}$. It is common [44] to define the error of the prediction at point p using logarithmic units as

$$E^s_p = P_{mes_p} - P^s_{pred_p} \tag{4.11}$$

being $ME^s = \overline{E^s_p}$ the mean error of the solution $\Omega^s_{N_m}$ and considering all N_p points, which can also be interpreted as the offset between the measurements and the predictions.

The root mean square error (RMSE), is also a common metric of the accuracy of a propagation model. It has thus been selected as the error function to be subject to the minimisation process. Then, for a given solution $\Omega^s_{N_m}$, the RMSE is calculated as

$$\mathrm{RMSE}(\Omega^s_{N_m}) = \sqrt{\frac{1}{N_p} \sum_{p=0}^{N_p-1} |E^s_p|^2}. \tag{4.12}$$

The objective of the calibration process is thus the minimisation of the RMSE subject to $ME = 0$.

[1] A solution is defined here as the set of numerical conductivities of the simulation materials.

4.2.1.2 Metaheuristics-based calibration

Once the error function has been defined, a brute-force approach to find an optimal solution to the problem could be used to test all possible solutions $\Omega^s_{N_m}$ and then choose the one that minimises the cost function (4.12). However, since the electrical conductivity takes values in the set of the real numbers, the space of solutions for $\Omega^s_{N_m}$ is infinite and a smarter approach is needed. Hence, a metaheuristics optimisation algorithm is proposed as a feasible way of searching the space of solutions. The algorithm used here is simulated annealing (SA), although the same concept also applies to other metaheuristic algorithms such as Taboo Search, Genetic Algorithm, etc.

SA [58] is an optimisation algorithm based on the physical technique of annealing, widely used in metallurgy. From the point of view of the minimisation of an error function for a given system, SA works by setting first the state of the system to a solution $\Omega^s_{N_m}$. Then, neighboring solutions $\Omega^{s'}_{N_m}$ are tested and the error function is evaluated in order to try to find a better solution, i.e. one such that $\text{RMSE}(\Omega^{s'}_{N_m}) < \text{RMSE}(\Omega^s_{N_m})$. If a better solution is found, then the current state of the system is updated to the new solution $\Omega^{s'}_{N_m}$. However, if a worst solution is found, the state of the system is set to this new neighbour solution with probability P. P is also called the acceptance probability function (APF) and it is a real function of $\text{RMSE}(\Omega^s_{N_m})$, $\text{RMSE}(\Omega^{s'}_{N_m})$ and a variable T called the temperature that is progressively decreased as the calibration progresses. When the temperature is high (i.e. at the beginning of the calibration process), worst solutions are very likely to be accepted (P close to one). However, when the temperature is low (i.e. towards the end of the calibration process), the acceptance probability of worst solutions decreases. According to these requirements, an appropriate APF has been designed. This is illustrated on Figure 4.3 and it is defined as

$$P(\Omega^s_{N_m}, \Omega^{s'}_{N_m}, T) = e^{\frac{RMSE(\Omega^s_{N_m}) - RMSE(\Omega^{s'}_{N_m})}{T}}. \tag{4.13}$$

Nevertheless, it must be remarked that the SA algorithm does not specify an APF. Thus, other calibration approaches may use a different one to their convenience.

The way the temperature T decreases between iterations of the calibration process is also subject to various implementations. For similarity with the physical

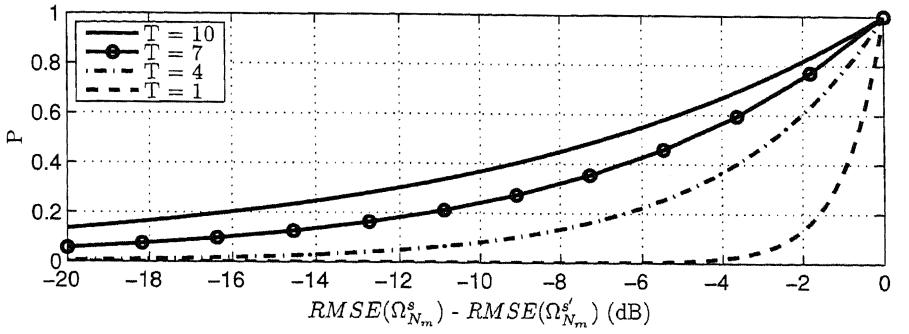


Figure 4.3: Acceptance probability function of equation (4.13) for the calibration process based on simulated annealing.

Table 4.5: Parameters of simulated annealing for calibration

Parameter	Value
AF	0.8286
T_0	10
L_T	50
N_T	10

annealing process, an exponential decrease of T has been used. Then, the value of the temperature at each iteration k of the calibration process is obtained with

$$T_k = AF * T_{k-1} \tag{4.14}$$

where $k \in [1, L_T - 1]$ and L_T is the number of different temperature levels. AF is called the annealing factor and it is related to the rate at which the temperature decreases from one iteration to the next. The calibration procedure starts at an initial temperature T_0, which is then decreased L_T times according to (4.14). On each temperature level, the system tests N_T different neighboring solutions, thus giving rise to a total of $N_{cal} = L_T \cdot N_T$ solutions being tested. The idea behind this, is to perform a deeper search on each temperature level before decreasing the chances of escaping local minima.

The way neighboring solutions are chosen in SA is also not specified by the algorithm ant it depends on the specifics of each application. Since the purpose here is to find the optimal values of the numerical electrical conductivity for a given set

of materials, only one material is modified at each stage. This guarantees a local search in the very neighbourhood of the current state.

Each of the solutions tested during calibration requires the evaluation of the cost function of equation (4.12), which implies the performance of an FDTD simulation. This is thus the most time-consuming part of the calibration process, whose total number of iterations N_{cal} must be considered in order to obtain results within an adequate time period. In general, N_{cal} depends not only on the number of materials involved in the simulation, but also on their accumulated presence throughout the scenario. For example, if 0.5% of all cells represented plaster and 99.5% represented air, finding a finer value of $\tilde{\sigma}_{plaster}$ would not change significantly the resulting power distribution. Hence, an increased number of calibration iterations to better optimise $\tilde{\sigma}_{plaster}$ looses significance. Considering then the execution time of a coverage simulation with $N_{\lambda_0} = 10$ at $f_0 = 3.5$ GHz and in the simulation sites described in section 4.1.2, $N_{cal} = 500$ calibration iterations have been performed. Previous calibrations using a substantially larger number of iterations ($N_{cal} = 2250$) were also performed. However, the resulting solutions were always the same as with $N_{cal} = 500$. When calibrations with $N_{cal} < 500$ were performed, the same solution was eventually found but some calibrations converged to worst solutions. Thus, $N_{cal} \approx 500$ represents a good compromise between the execution time of the calibration and the goodness of the obtained results. The remaining parameters of the calibration procedure are summarised on Table 4.5.

4.2.2 Small-scale fading removal filter

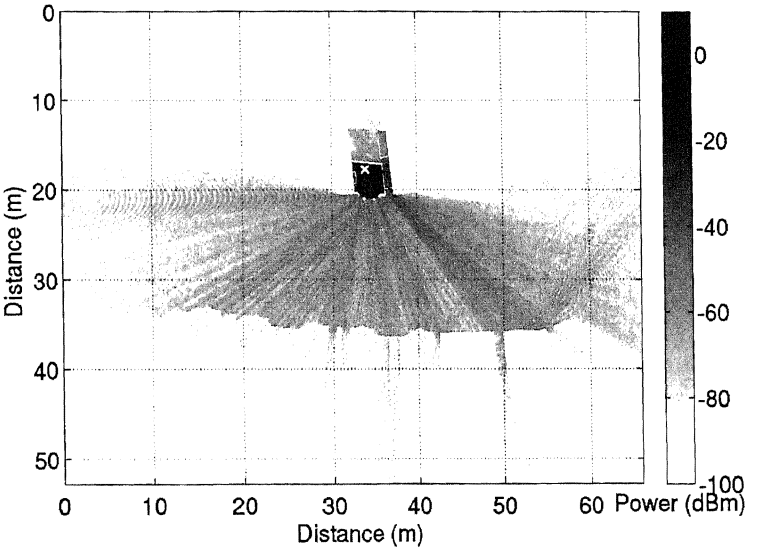
In a similar way to what happens with CW power measurements, narrowband FDTD simulations are also subject to small-scale fading. This is due to the fact that FDTD describes the propagation of numerical waves, which interfere each other in a similar way to the physical phenomenon of wave interference. Since the objective is to model the path loss, fading can be removed from these simulations using a spatial filter similar to the one described in section 4.1.4 for the processing of power measurements. However, the output of an FDTD simulation contains samples at a very fine spatial resolution across the simulation site. Thus, contrarily to the scattered nature of empirical power measurements, numerical simulations provide

a large quantity of uniformly distributed power samples. Therefore, using a large spatial filter of radius $40 \cdot \lambda_0$ as in section 4.1.4 is not necessary, as it would use up an unnecessarily large amount of power samples. Instead, a filter of size $4\lambda_0 \times 4\lambda_0$ is enough to average most of the effects that take place over distances on the order of a wavelength (e.g. small-scale fading). The size of such a filter is 576 cells when $N_{\lambda_0} = 6$ and 1600 cells when $N_{\lambda_0} = 10$, which suffices for the estimation of the power average. The effect of such a fading removal filter is illustrated on Figure 4.4.

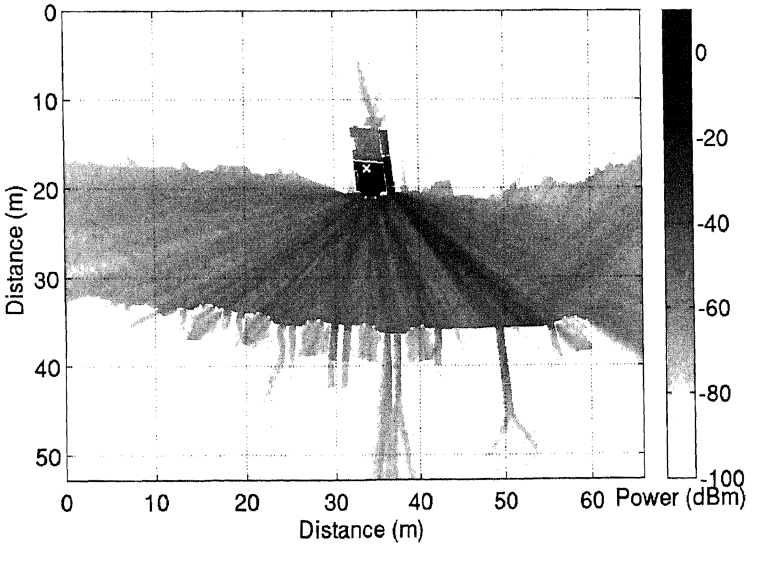
4.2.3 Accuracy validation

The measurements-based calibration procedure can also be seen as the interpolation of the output of a system with several degrees of freedom. In this case, such system is the FDTD algorithm and the parameters to tune are the numerical electrical conductivities of the materials involved in the simulation. In addition, this calibration process could also be applied to lower-frequency FDTD simulations. This is particularly interesting because the simulation running time would be reduced as a consequence of the smaller grid size. However, although the path loss can be well modelled through calibration, the phase behaviour and the fading patterns vary at different frequencies. Furthermore, it was the purpose of section 3.2 to highlight that the behaviour of diffractive phenomena varies with the simulation frequency. Thus, in order to assess the accuracy with which coverage predictions are performed at the original (FRF $= 1$) and lower (FRF < 1) frequencies, several calibrations at different frequency reduction factors (FRFs) have been performed. Furthermore, to analyse the global validity of the calibrated conductivities, results obtained from two different simulation sites are presented. First, the calibration is performed on one site and then the calibrated material parameters are reused for simulation on a different site. The simulation sites under consideration here are those described in section 4.1.

To present the results, Tables 4.6 to 4.8 contain the RMSE values resulting from the calibration process on Site 1. These Tables also present the RMSE of simulations performed on Site 2 using the material parameters resulting from the calibration on Site 1. The shaded columns highlight the results that correspond to simulations



(a) Before small-scale fading removal.



(b) After small-scale fading removal.

Figure 4.4: Coverage prediction with $N_{\lambda_0} = 10$ at $f_0 = 3.5$ GHz and 10 dBm transmitter power.

Table 4.6: RMSE. Calibration on Site 1. $f = 0.9$ GHz

Site	N_{λ_0}	FRF					
		0.031	0.062	0.124	0.249	0.499	1.000
1	6	–	–	–	13.15	5.22	4.17
	10	–	–	3.64	7.45	6.21	2.94
2	6	–	–	–	5.37	6.58	4.91
	10	–	–	6.31	7.03	6.53	4.23

Table 4.7: RMSE. Calibration on Site 1. $f = 2.5$ GHz

Site	N_{λ_0}	FRF					
		0.031	0.062	0.124	0.249	0.499	1.000
1	6	–	–	7.25	7.1	5.45	5.92
	10	–	8.64	6.04	6.07	6.56	5.78
2	6	–	–	7.99	8.17	7.53	5.28
	10	–	7.69	7.89	6.84	6.37	5.26

at the original frequency (FRF = 1). Then, Tables 4.9 to 4.11 contain the RMSE resulting from the calibration process on Site 2, as well as the RMSE evaluated on Site 1. These results reveal that the RMSE on one site does not necessarily match the RMSE on the other. This difference in the RMSE across sites has an average absolute value of 1.6 dB and, from these results, it takes negative and positive values with equal probability. Hence, the error on the calibrated site is not always lower than on the evaluation site, thus increasing the reliability of the calibration method. Finally, the RMSE at the original frequency and calculated across all frequencies and sites is 5.2 dB.

Reducing the simulation frequency decreases also the spatial step. Therefore, the precision with which the simulation scenario can be described diminishes for lower FRF values. In order to avoid errors in the numerical simulations due to a coarse spatial discretisation, a minimum spatial step at which to simulate has been imposed. This is chosen in terms of the smallest architectural detail modelled in the current simulation sites, which corresponds to a small window in one of the houses that is 30 centimeter wide. Therefore, in Tables 4.6 to 4.17 no results are shown for FRF values corresponding to spatial steps smaller than 30 centimeter.

Table 4.8: RMSE. Calibration on Site 1. $f = 3.5$ GHz

Site	N_{λ_0}	FRF					
		0.031	0.062	0.124	0.249	0.499	1.000
1	6	–	5.27	6.8	7.32	7.37	5.62
	10	3.66	4.68	8.05	8.06	6.97	5.51
2	6	–	7.44	8.11	8.54	6.95	2.11
	10	8.83	6.75	7.35	5.9	6.22	5.77

Table 4.9: RMSE. Calibration on Site 2. $f = 0.9$ GHz

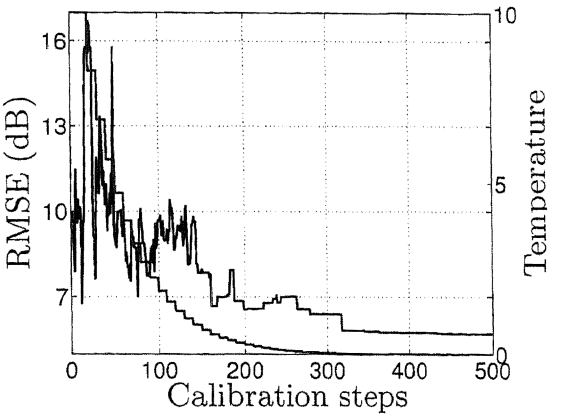
Site	N_{λ_0}	FRF					
		0.031	0.062	0.124	0.249	0.499	1.000
1	6	–	–	–	7.72	2.8	5.71
	10	–	–	3.15	6.91	3.4	4.17
2	6	–	–	–	6.14	5.13	4.93
	10	–	–	6.04	4.9	5.13	4.88

Table 4.10: RMSE. Calibration on Site 2. $f = 2.5$ GHz

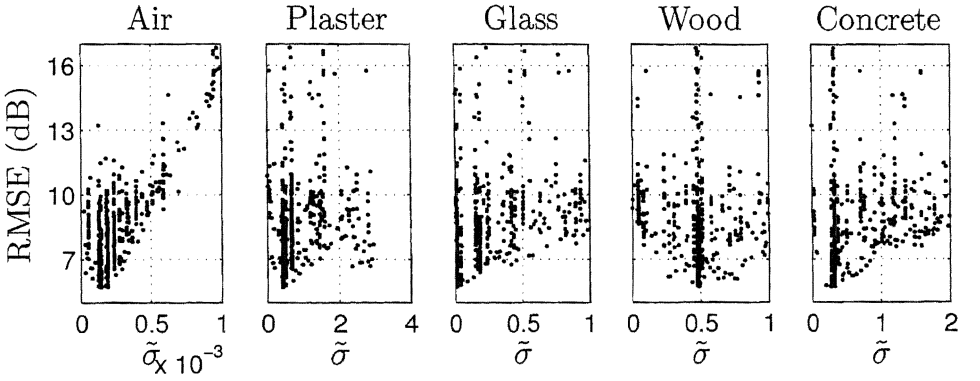
Site	N_{λ_0}	FRF					
		0.031	0.062	0.124	0.249	0.499	1.000
1	6	–	–	7.31	7.11	7.57	5.50
	10	–	11.44	5.97	6.59	6.96	5.63
2	6	–	–	6.91	6.66	5.71	5.10
	10	–	5.95	6.66	6.84	5.98	5.71

Table 4.11: RMSE. Calibration on Site 2. $f = 3.5$ GHz

Site	N_{λ_0}	FRF					
		0.031	0.062	0.124	0.249	0.499	1.000
1	6	–	7.75	6.13	6.17	7.31	6.20
	10	3.86	5.90	7.58	8.64	6.49	5.52
2	6	–	9.44	7.5	7.32	7.14	5.65
	10	6.17	8.07	7.38	7.7	7.60	5.41



(a) Evolution of the cost function through calibration.



(b) Samples of the space of solutions.

Figure 4.5: Example of calibration process on Site 2 with $N_{\lambda_0} = 10$ at $f_0 = 2.5$ GHz and $FRF = 1$.

From the RMSE results, it is also observed that $N_{\lambda_0} = 10$ only performs, on average, slightly better than $N_{\lambda_0} = 6$. Thus, $N_{\lambda_0} = 6$ may be a wiser choice for the performance of path-loss predictions that are less computationally expensive. However and as will be seen in chapter 6, the grid resolution requirements are stronger when wideband simulations are considered.

4.2.4 Outcome of the calibration process

As an example of the sounding of the space of solutions that the calibration process performs, Figure 4.5(a) shows the evolution of the cost function as the temperature decreases. This illustrates how, when the temperature is high, the system eventually

Table 4.12: Calibrated electrical conductivity $\tilde{\sigma}$ on Site 1. $f = 0.9$ GHz

N_{λ_0}	Material	FRF					
		0.031	0.062	0.124	0.249	0.499	1.000
6	Air	–	–	–	1.0E-4	9.5E-5	9.8E-5
	Plaster	–	–	–	3.1E-1	3.8E-1	2.9E0
	Glass	–	–	–	9.9E-1	4.7E-2	2.8E-1
	Wood	–	–	–	1.0E-1	6.8E-1	9.9E-1
	Concrete	–	–	–	1.3E0	1.1E0	2.0E0
10	Air	–	–	9.7E-5	9.6E-5	8.6E-5	3.4E-4
	Plaster	–	–	2.8E0	1.0E0	1.8E0	4.0E-1
	Glass	–	–	2.6E-2	7.4E-1	2.1E-1	7.8E-1
	Wood	–	–	1.0E-3	2.3E-1	2.1E-1	2.3E-1
	Concrete	–	–	1.8E0	1.6E-1	3.4E-1	6.5E-1

selects worst solutions to escape local minima. However, as T decreases, only better solutions are allowed. Furthermore, Figure 4.5(b) illustrates the behaviour of the cost function across the five-dimensional space of solutions. This shows that some $\tilde{\sigma}$ values perform substantially better than others (lower RMSE), thus highlighting the need for the calibration process.

Finally, Tables 4.12 to 4.14 contain the obtained values of the numerical electrical conductivity when the calibration is performed on Site 1. Similarly the values obtained when calibrating on Site 2 are displayed on Tables 4.15 to 4.17. Further, it is observed from these that the conductivity values from different sites do not, in general, differ more than one order of magnitude. It must also be remarked that, since cells characterising air represent approximately 92% of the whole scenario, the uncertainty in the value of the numerical electrical conductivity of air is responsible for most of the simulation error. Furthermore, $\tilde{\sigma}_{air}$ is responsible for the power decay with distance. It is hence this parameter that hinders the non-quadratic spatial decay of cylindrical waves in 2D FDTD simulations.

Table 4.13: Calibrated electrical conductivity $\tilde{\sigma}$ on Site 1. $f = 2.5$ GHz

N_{λ_0}	Material	FRF					
		0.031	0.062	0.124	0.249	0.499	1.000
6	Air	–	–	7.9E-5	8.7E-5	5.4E-7	3.1E-4
	Plaster	–	–	2.9E0	8.1E-1	6.5E-2	1.3E0
	Glass	–	–	2.4E0	7.1E-1	9.2E-1	3.4E-1
	Wood	–	–	1.7E0	9.7E-1	2.7E-1	9.9E-1
	Concrete	–	–	1.9E0	1.3E0	2.6E-1	2.0E0
10	Air	–	9.8E-5	2.6E-5	7.1E-5	3.7E-5	5.4E-5
	Plaster	–	4.2E-1	3.7E-2	3.8E-1	3.6E-1	5.1E-1
	Glass	–	7.5E-1	2.2E-1	2.2E-1	1.7E-2	5.9E-1
	Wood	–	4.2E-4	4.9E-1	3.2E-1	3.8E-1	9.5E-2
	Concrete	–	8.6E-2	7.1E-2	1.5E-1	3.8E-2	8.8E-1

Table 4.14: Calibrated electrical conductivity $\tilde{\sigma}$ on Site 1. $f = 3.5$ GHz

N_{λ_0}	Material	FRF					
		0.031	0.062	0.124	0.249	0.499	1.000
6	Air	–	9.5E-5	7.8E-6	9.6E-5	9.6E-5	9.7E-5
	Plaster	–	1.4E-1	2.9E0	4.6E-1	3.1E-1	1.3E0
	Glass	–	9.5E-1	1.3E-1	2.9E-1	7.1E-1	9.6E-1
	Wood	–	5.0E-2	8.8E-1	4.0E-1	7.9E-1	1.3E-1
	Concrete	–	5.7E-3	2.0E0	1.8E-2	8.0E-1	2.9E-1
10	Air	7.2E-5	1.0E-4	9.8E-5	9.4E-5	9.7E-5	3.8E-4
	Plaster	2.9E0	3.1E-1	5.3E-2	2.3E-3	2.7E0	6.5E-2
	Glass	6.1E-1	3.1E-1	3.9E-1	3.8E-1	9.8E-1	7.6E-1
	Wood	7.2E-2	8.1E-2	1.7E-2	6.8E-1	3.0E-1	1.8E-1
	Concrete	4.7E-2	1.5E0	7.8E-1	9.1E-2	1.7E-2	1.1E0

Table 4.15: Calibrated electrical conductivity $\tilde{\sigma}$ on Site 2. $f = 0.9$ GHz

N_{λ_0}	Material	FRF					
		0.031	0.062	0.124	0.249	0.499	1.000
6	Air	–	–	–	7.7E-5	4.1E-4	2.0E-4
	Plaster	–	–	–	2.7E0	3.4E-1	8.3E-1
	Glass	–	–	–	9.9E-2	1.7E-2	6.8E-1
	Wood	–	–	–	9.7E-1	8.1E-1	2.4E-1
	Concrete	–	–	–	5.6E-2	3.8E-1	1.8E0
10	Air	–	–	9E-5	6.0E-4	2.2E-4	2.7E-4
	Plaster	–	–	2.0E-2	2.0E-2	3.8E-1	2.8E0
	Glass	–	–	1.2E-1	1.2E-1	4.8E-3	9.8E-1
	Wood	–	–	3.8E-1	3.8E-1	2.7E-1	2.3E-1
	Concrete	–	–	1.3E-1	1.3E-1	2.7E-2	1.3E-1

Table 4.16: Calibrated electrical conductivity $\tilde{\sigma}$ on Site 2. $f = 2.5$ GHz

N_{λ_0}	Material	FRF					
		0.031	0.062	0.124	0.249	0.499	1.000
6	Air	–	–	2.0E-4	1.6E-4	3.7E-4	1.9E-4
	Plaster	–	–	4.7E-1	7.1E-1	4.6E-2	9.5E-1
	Glass	–	–	1.2E-1	1.4E-1	3.6E0	2.2E-1
	Wood	–	–	4.1E-1	9.3E-2	6.3E-1	3.2E-1
	Concrete	–	–	9.3E-2	2.2E-1	5.0E-2	1.9E-2
10	Air	–	4.1E-4	2.2E-4	2.2E-4	2.1E-4	1.4E-4
	Plaster	–	9.2E-2	4.2E-1	2.0E0	2.2E0	4.2E-1
	Glass	–	4.6E-1	5.1E-2	1.9E-2	4.0E-1	1.3E-2
	Wood	–	3.7E-1	9.3E-1	4.7E-2	1.7E-1	4.9E-1
	Concrete	–	7.0E-2	2.4E-1	1.8E0	4.4E-2	2.8E-1

Table 4.17: Calibrated electrical conductivity $\tilde{\sigma}$ on Site 2. $f = 3.5$ GHz

N_{λ_0}	Material	FRF					
		0.031	0.062	0.124	0.249	0.499	1.000
6	Air	–	3.6E-4	2.6E-4	3.2E-4	1.6E-4	9.3E-7
	Plaster	–	1.3E-1	3.7E-1	2.9E0	3.0E0	1.3E-1
	Glass	–	2.2E-1	8.9E-1	9.3E-1	2.2E-1	3.0E-3
	Wood	–	2.1E-1	1.9E-2	6.4E-2	9.9E-1	1.3E-2
	Concrete	–	1.2E-1	1.8E-2	1.2E0	4.8E-1	3.6E-1
10	Air	3.1E-4	1.1E-5	1.5E-4	4.5E-5	3.3E-4	1.4E-4
	Plaster	1.9E-1	1.7E-1	2.6E0	3.9E-1	1.1E-1	1.4E0
	Glass	4.0E-1	1.0E-1	6.1E-2	9.1E-1	9.9E-1	1.1E-1
	Wood	1.9E-1	5.9E-1	8.0E-1	3.2E-1	5.1E-1	1.0E0
	Concrete	7.2E-3	2.6E-1	1.1E0	5.4E-2	1.9E0	1.1E0

4.3 Conclusions

This chapter has dealt with a measurements-based methodology for the calibration of the materials involved in coverage predictions with FDTD. The purpose of this model tuning approach, based on the metaheuristics of simulated annealing, is to map the propagation predictions to the reality. Then, the method has been validated by reusing the calibrated material properties for the coverage prediction of a different but similar simulation site. Finally, calibration results were presented in both scenarios at different RF frequencies and at different frequency reduction factors.

4.3.1 Contributions

To summarise, the main contributions of this chapter made by the author to the scientific community are described in the following. Furthermore, these are illustrated through concrete numerical examples:

1. **Novel calibration approach based on simulated annealing (SA) for FDTD:** The calibration procedure described in this chapter adds on to the existing literature on the use of metaheuristics for the calibration of radio

propagation models. Further, it shows that SA is an appropriate tool for this purpose and that it finds near-optimal solutions of the problem under consideration in some hundred iterations.

To illustrate the calibration speed, let us recall the example used in section 3.4.1, which considered a scenario of size 100×50 meters to be calibrated at $f_{sim} = 900$ MHz. Let us also use a sampling step ten times smaller than the physical wavelength ($N_{\lambda_{sim}} = 10$) and a Courant number of $S = 0.9$. Furthermore, let us perform the calibration simulations in the TESLA C870 card described in Table 3.1. Using then the implementation described in the previous chapter, a computational speed of approximately 800 Mcps is reached. Then, 500 calibration iterations can be performed in about one hour and 18 minutes. As can be seen in Tables 4.6 and 4.9, this calibration technique improves the prediction accuracy by bringing the RMSE to values well below 6 dB.

2. **Characterisation of path loss in indoor-to-outdoor residential scenarios:** The power measurements obtained during this thesis describe the coverage area of femtocells over a wide range of frequencies. Furthermore, the material parameters obtained from the calibration process using these measurements allow any individual to further recreate and study the femtocell radio channel.

   For instance, Figure 4.4 showed that house-to-house interference is, in general, negligible. However, it is also seen that the street immediately adjacent to the transmitter premises suffers from large interference levels. Finally, the numerical conductivity values of Tables 4.12 to 4.17 can be directly reused in similar simulation scenarios by third parties.

The importance of reliable coverage prediction methods for the planning and optimisation of wireless networks is well known [2]. However, these are not enough to study the additional distortions that the propagation channel causes to wideband systems. For instance, future LTE-Advanced systems with bandwidths of up to 100 MHz will be subject to frequency selective fading. The modelling of such phenomena is usually based on CIR predictions, which could be obtained from FDTD simulations of pulsed sources. However, the use of FDTD for wideband simulations

in large areas is not yet consolidated due to the large computational requirements. Furthermore, the inherent dispersive properties of the FDTD grid introduce errors in the simulation of wideband pulsed sources that are not fully understood. Hence the next chapter investigates in depth and achieves to explain two of the main problems of this type of simulations, thus opening the door for error-bounded wideband FDTD simulations of low computational load.

5

Pulse propagation with the finite-difference time-domain method

Yee grids have numerical dispersive properties that lead to the anisotropy of the numerical phase velocity, which is a source of error in CW simulations. As a result, the behaviour of the grid is frequency dependant, which affects the simulation of sources with energy content over wide spectral regions. Within this context, the main contribution of this chapter is the introduction of several analytical formulae for the quantification of errors in the simulation of wideband pulses. These errors are:

- The spatial anisotropy of the numerical group velocity.
- The time spread of numerical pulses.

Wideband FDTD simulations in large scale scenarios have been performed already in the past. In particular for the purpose of UWB [59] [60] and channel impulse response (CIR) prediction [22]. However, these studies have normally used large spatial sampling resolutions ($N_{\lambda_0} \geq 30$), thus giving rise to huge computational requirements. Most of these simulations take weeks to run and consume large amounts of computer memory. Fortunately, the effects of numerical dispersion are negligible at such large spatial resolutions and are usually unaccounted for.

In the interest of performance in the computation of wideband FDTD simulations, large spatial and time steps are preferred. However in this case, the effects of numerical dispersion can not be disregarded. If the accuracy requirements for a given prediction were known a priori, it would then be possible to configure the simulation such that errors due to numerical dispersion would be bounded to the desired limits while also reducing the computational load. In this chapter, the above described sources of error have been researched and quantified. As a result, the formulae developed here describe fundamental properties of the standard Yee algorithm. From a practical standpoint, these results open the door to the design of FDTD grids for wideband simulation with bounded numerical dispersion errors and large spatial and time steps.

5.1 Fundamentals

Let Δx, Δy and Δz be the cell sizes along the x, y and z dimensions in a 3D staggered Yee grid, and let the cell aspect ratios be those defined by equations (2.15) and (2.16). Further, let the numerical dispersion relationship be that of equation (2.14) which, for readability, is expanded next as

$$\begin{aligned}\frac{1}{(\Delta x)^2}sin^2\left(\frac{\tilde{k}_x\Delta x}{2}\right)+\frac{1}{(\Delta y)^2}sin^2\left(\frac{\tilde{k}_y\Delta y}{2}\right)+\frac{1}{(\Delta z)^2}sin^2\left(\frac{\tilde{k}_z\Delta z}{2}\right)\\ =\frac{1}{(c\Delta t)^2}sin^2\left(\frac{\omega\Delta t}{2}\right)\end{aligned} \tag{5.1}$$

where c is the speed of light in the medium being modelled, Δt the time step and ω the angular frequency. Let also ϕ and θ be respectively the azimuth and elevation angles of the propagation direction. Then, the Cartesian projections ($\tilde{k}_x$, $\tilde{k}_y$ and $\tilde{k}_z$) of the numerical wave vector and its module $\tilde{k}$ are

$$\tilde{k}_x = \tilde{k}sin(\theta)cos(\phi) \tag{5.2}$$

$$\tilde{k}_y = \tilde{k}sin(\theta)sin(\phi) \tag{5.3}$$

$$\tilde{k}_z = \tilde{k}cos(\theta) \tag{5.4}$$

$$\tilde{k} = \sqrt{\tilde{k}_x^2 + \tilde{k}_y^2 + \tilde{k}_z^2}. \tag{5.5}$$

Finally, let the Courant number be the one defined by equation (2.18), which is also expanded next for readability:

$$S = \sqrt{\left(\frac{c\Delta t}{\Delta x}\right)^2 + \left(\frac{c\Delta t}{\Delta y}\right)^2 + \left(\frac{c\Delta t}{\Delta z}\right)^2}. \tag{5.6}$$

5.2 Optimisation of the numerical group velocity

As mentioned above, FDTD simulations have already been used for the study of the propagation conditions that wideband radio communication systems are subject to. However, it must be remarked that the propagation characteristics of wavepackets and that of monochromatic signals in an FDTD grid are fundamentally different. On the one hand single-frequency fields propagate through the grid at a speed given by the numerical phase velocity $\tilde{v}_p$. On the other hand, the speed of propagation of pulsed signals is governed by the numerical group velocity $\tilde{v}_g$. Since the dispersion relationship of Yee's grid is not linear, these velocities are not necessarily equal and they vary along different spatial directions. As a result, numerical electric and magnetic wavepackets propagate with different speeds depending on the propagation direction under consideration. The anisotropy suffered by the phase and group velocity in Yee's FDTD is thus a fundamental source of error in computational electromagnetic simulations.

The extrema of the numerical phase velocity $\tilde{v}_p$ were found by Zhao [32] in 2003, thus also characterising its anisotropy. However, the numerical group velocity $\tilde{v}_g$ has received little attention from the scientific community. This is odd, as the quantification of the anisotropy error suffered by $\tilde{v}_g$ is an essential tool for the evaluation of the accuracy achieved by pulsed-source simulations based on FDTD. If this error is too high, it can lead to incorrect estimations of the time of arrival (ToA) of multipath components (MCs) in simulations with pulsed-sources. The anisotropy error of $\tilde{v}_g$ is defined by its extrema throughout the grid and hence, their knowledge is fundamental for its characterisation.

In order to quantify uncertainties in the propagation of wavepackets due to the anisotropy of $\tilde{v}_g$, the numerical group velocity anisotropy error is defined as

$$\Delta\tilde{v}_{g_{aniso}} = \frac{\hat{v}_M - \hat{v}_m}{\hat{v}_m} \tag{5.7}$$

where $\hat{v}_M = max\{\tilde{v}_g\}$ and $\hat{v}_m = min\{\tilde{v}_g\}$. To evaluate this expression, it is thus necessary to obtain first the extrema of the numerical group velocity, i.e. to optimise $\tilde{v}_g$. This was attempted in 2007 [57] but, due to errors in the derivation process, the maximum was not found. Instead, a conservative upper bound to $\hat{v}_M$ was provided. In this section, the problem of the optimisation of $\tilde{v}_g$ is reformulated and the resulting equations solved rigorously using the method of Lagrange multipliers. This way, tight values for the minimum and maximum of the numerical group velocity are obtained. Based on this, (5.7) can be fully characterised, thus assessing the errors to which the propagation of numerical wavepackets in staggered FDTD grids are subject to.

5.2.1 Extrema in 3D

The overall shape of an electromagnetic wave moves at the group velocity, which is defined as $v_g \equiv \partial\omega/\partial k$, where k is the wave number[1]. In a Yee grid, the numerical wave number $\tilde{k}$ varies with the angular direction, hence causing the anisotropy of the numerical phase and group velocities. Therefore, by clearing ω in the dispersion relationship of (5.1) and differentiating with respect to $\tilde{k}$, the numerical group velocity $\tilde{v}_g$ in 3D FDTD can be expressed as

$$\tilde{v}_g = A \cdot \left(\frac{\tilde{k}_x sin(\tilde{k}_x \Delta x)}{\tilde{k}\Delta x} + \frac{\tilde{k}_y sin(\tilde{k}_y \Delta y)}{\tilde{k}\Delta y} + \frac{\tilde{k}_z sin(\tilde{k}_z \Delta z)}{\tilde{k}\Delta z} \right) \tag{5.8}$$

where the auxiliary variable is

$$A = \frac{c^2}{\omega sinc(\omega\Delta t)} \tag{5.9}$$

and $sinc(x) = sin(x)/x$ is the unnormalised sinc function. Due to the symmetry of the FDTD grid, the following analysis is presented only in the interval $\phi \in [0, \pi/2]$ and $\theta \in [0, \pi/2]$. The results can thus be extrapolated to other directions by simply rotating the grid.

[1]The wave number is the module of the wave vector.

Theorem 1. *The maximum and minimum values of the numerical group velocity in standard 3D FDTD are*

$$\hat{v}_M = \frac{c}{cos(\frac{\omega\Delta t}{2})}\sqrt{1-\frac{1}{S^2}sin^2\left(\frac{\omega\Delta t}{2}\right)} \tag{5.10}$$

$$\hat{v}_m = \min_{\zeta\in\{x,y,z\}}\left\{\frac{c}{cos(\frac{\omega\Delta t}{2})}\sqrt{1-\left(\frac{\Delta\zeta}{c\Delta t}\right)^2 sin^2\left(\frac{\omega\Delta t}{2}\right)}\right\} \tag{5.11}$$

being the maximum achieved for the propagation direction

$$\phi_{\hat{v}_M} = arctan(R_y) \tag{5.12}$$

$$\theta_{\hat{v}_M} = arctan\left(\frac{\sqrt{1+R_y^2}}{R_z}\right) \tag{5.13}$$

and being the minimum achieved for the propagation direction

$$\phi_{\hat{v}_m} = \begin{cases} 0, & if\ x=\zeta_M \\ \frac{\pi}{2}, & if\ y=\zeta_M \\ [0,\pi/2], & if\ z=\zeta_M \end{cases} \tag{5.14}$$

$$\theta_{\hat{v}_m} = \begin{cases} 0, & if\ z=\zeta_M \\ \frac{\pi}{2}, & if\ x,y=\zeta_M \end{cases} \tag{5.15}$$

where

$$\zeta_M = \underset{\zeta\in\{x,y,z\}}{argmax}\{\Delta\zeta\}. \tag{5.16}$$

Proof. Finding the extrema of $\tilde{v}_g$ relies on optimising (5.8) under the constraint given by (5.1). Formulating this problem using Lagrange multipliers leads to the following Lagrangian:

$$\begin{aligned}\mathcal{L}_g(\tilde{k}_x,\tilde{k}_y,\tilde{k}_z,\eta_g) = {} & \frac{\eta_g}{(c\Delta t)^2}sin^2\left(\frac{\omega\Delta t}{2}\right)+ \\ & \left[\frac{A}{\Delta x}\frac{\tilde{k}_x sin(\tilde{k}_x\Delta x)}{\tilde{k}}-\frac{\eta_g}{(\Delta x)^2}sin^2\left(\frac{\tilde{k}_x\Delta x}{2}\right)\right]+ \\ & \left[\frac{A}{\Delta y}\frac{\tilde{k}_y sin(\tilde{k}_y\Delta y)}{\tilde{k}}-\frac{\eta_g}{(\Delta y)^2}sin^2\left(\frac{\tilde{k}_y\Delta y}{2}\right)\right]+ \\ & \left[\frac{A}{\Delta z}\frac{\tilde{k}_z sin(\tilde{k}_z\Delta z)}{\tilde{k}}-\frac{\eta_g}{(\Delta z)^2}sin^2\left(\frac{\tilde{k}_z\Delta z}{2}\right)\right]\end{aligned} \tag{5.17}$$

where η_g is a Lagrange multiplier determined by the solution to the system that arises from differentiating (5.17) and equalling to zero. Differentiating (5.17) with respect to η_g and equalling to zero returns the dispersion relationship (5.1). Then, the other partial derivatives are

$$\begin{aligned}\frac{\partial \mathcal{L}_g}{\partial \tilde{k}_x} =& \frac{A}{\Delta x}\frac{\tilde{k}^2 - \tilde{k}_x^2}{\tilde{k}^3} sin(\tilde{k}_x \Delta x) + A\frac{\tilde{k}_x}{\tilde{k}} cos(\tilde{k}_x \Delta x) \\ &- \frac{A}{\Delta y}\frac{\tilde{k}_x \tilde{k}_y}{\tilde{k}^3} sin(\tilde{k}_y \Delta y) - \frac{A}{\Delta z}\frac{\tilde{k}_x \tilde{k}_z}{\tilde{k}^3} sin(\tilde{k}_z \Delta z) \\ &- \frac{\eta_g}{2\Delta x} sin(\tilde{k}_x \Delta x)\end{aligned} \tag{5.18}$$

$$\begin{aligned}\frac{\partial \mathcal{L}_g}{\partial \tilde{k}_y} =& \frac{A}{\Delta y}\frac{\tilde{k}^2 - \tilde{k}_y^2}{\tilde{k}^3} sin(\tilde{k}_y \Delta y) + A\frac{\tilde{k}_y}{\tilde{k}} cos(\tilde{k}_y \Delta y) \\ &- \frac{A}{\Delta x}\frac{\tilde{k}_y \tilde{k}_x}{\tilde{k}^3} sin(\tilde{k}_x \Delta x) - \frac{A}{\Delta z}\frac{\tilde{k}_y \tilde{k}_z}{\tilde{k}^3} sin(\tilde{k}_z \Delta z) \\ &- \frac{\eta_g}{2\Delta y} sin(\tilde{k}_y \Delta y)\end{aligned} \tag{5.19}$$

$$\begin{aligned}\frac{\partial \mathcal{L}_g}{\partial \tilde{k}_z} =& \frac{A}{\Delta z x}\frac{\tilde{k}^2 - \tilde{k}_z^2}{\tilde{k}^3} sin(\tilde{k}_z \Delta z) + A\frac{\tilde{k}_z}{\tilde{k}} cos(\tilde{k}_z \Delta z) \\ &- \frac{A}{\Delta x}\frac{\tilde{k}_z \tilde{k}_x}{\tilde{k}^3} sin(\tilde{k}_x \Delta x) - \frac{A}{\Delta y}\frac{\tilde{k}_z \tilde{k}_y}{\tilde{k}^3} sin(\tilde{k}_y \Delta y) \\ &- \frac{\eta_g}{2\Delta z} sin(\tilde{k}_z \Delta z).\end{aligned} \tag{5.20}$$

In the following analysis, let $\tilde{k}_\zeta$ and $\Delta\zeta$ with $\zeta \in \{x, y, z\}$ denote the different components of the wave vector and the spatial steps. Then, by equalling (5.18), (5.19) and (5.20) to zero, it can be seen that the extrema of $\tilde{v}_g$ in 3D occur under three different conditions:

Only one $\tilde{k}_\zeta$ is nonzero: Let $\tilde{k}_\nu$ with $\nu \in \{x, y, z\}$ be the nonzero wave vector component. In this case $\tilde{k}_\nu$ is directly obtained from (5.1). By replacing it in (5.8), the following extremum of the numerical group velocity is obtained:

$$\tilde{v}_g = \frac{c}{cos(\frac{\omega \Delta t}{2})}\sqrt{1 - \left(\frac{\Delta \nu}{c\Delta t}\right)^2 sin^2\left(\frac{\omega \Delta t}{2}\right)}. \tag{5.21}$$

Two $\tilde{k}_\zeta$ are nonzero: When only one component of the wave vector is zero, the problem is reduced to one of three possible 2D cases:

- If $\tilde{k}_z = 0$, equations (5.18), (5.19) and (5.20) reduce to

$$\begin{aligned}\frac{\partial \mathcal{L}_g}{\partial \tilde{k}_x} =& \frac{A}{\Delta x}\frac{\tilde{k}^2 - \tilde{k}_x^2}{\tilde{k}^3} sin(\tilde{k}_x \Delta x) + A\frac{\tilde{k}_x}{\tilde{k}} cos(\tilde{k}_x \Delta x) \\ &- \frac{A}{\Delta y}\frac{\tilde{k}_x \tilde{k}_y}{\tilde{k}^3} sin(\tilde{k}_y \Delta y) - \frac{\eta_g}{2\Delta x} sin(\tilde{k}_x \Delta x)\end{aligned} \tag{5.22}$$

$$\begin{aligned}\frac{\partial \mathcal{L}_g}{\partial \tilde{k}_y} =& \frac{A}{\Delta y}\frac{\tilde{k}^2 - \tilde{k}_y^2}{\tilde{k}^3} sin(\tilde{k}_y \Delta y) + A\frac{\tilde{k}_y}{\tilde{k}} cos(\tilde{k}_y \Delta y) \\ &- \frac{A}{\Delta x}\frac{\tilde{k}_y \tilde{k}_x}{\tilde{k}^3} sin(\tilde{k}_x \Delta x) - \frac{\eta_g}{2\Delta y} sin(\tilde{k}_y \Delta y).\end{aligned} \tag{5.23}$$

  The system that arises from equalling these to zero is solved in Appendix A, where it is seen that the unique solution is $\Delta x \cdot \tilde{k}_x = \Delta y \cdot \tilde{k}_y$, thus leading to the following extremum:

$$\tilde{v}_g = \frac{c}{cos(\frac{\omega \Delta t}{2})}\sqrt{1 - \left(\frac{\Delta x}{c\Delta t}\right)^2 \frac{1}{1+R_y^2} sin^2\left(\frac{\omega \Delta t}{2}\right)} \tag{5.24}$$

- If $\tilde{k}_y = 0$ and following the same 2D approach as in the previous case, the following extremum arises:

$$\tilde{v}_g = \frac{c}{cos(\frac{\omega \Delta t}{2})}\sqrt{1 - \left(\frac{\Delta x}{c\Delta t}\right)^2 \frac{1}{1+R_z^2} sin^2\left(\frac{\omega \Delta t}{2}\right)} \tag{5.25}$$

- Similarly, if $\tilde{k}_x = 0$ then

$$\tilde{v}_g = \frac{c}{cos(\frac{\omega \Delta t}{2})}\sqrt{1 - \left(\frac{\Delta x}{c\Delta t}\right)^2 \frac{1}{R_y^2+R_z^2} sin^2\left(\frac{\omega \Delta t}{2}\right)} \tag{5.26}$$

All $\tilde{k}_\zeta$ are nonzero: In this case, the system formed by equalling equations (5.18), (5.19) and (5.20) to zero must be solved, which is done in Appendix B. The unique solution is then found to be $\Delta x \tilde{k}_x = \Delta y \tilde{k}_y = \Delta z \tilde{k}_z$. This implies that this extremum of the numerical group velocity occurs in the same direction as the maximum of the numerical phase velocity (see [32]), being its value

$$\tilde{v}_g = \frac{c}{cos(\frac{\omega \Delta t}{2})}\sqrt{1 - \frac{1}{S^2} sin^2\left(\frac{\omega \Delta t}{2}\right)}. \tag{5.27}$$

Equations (5.21), (5.24), (5.25), (5.26) and (5.27) are the extrema of $\tilde{v}_g$ in 3D. However, it still remains to be determined which one is the maximum and which one is the minimum. A comparison is thus made in the following:

Comparing (5.27) **with** (5.21): If $\tilde{k}_\nu$ is the nonzero component in (5.21), and S is given by (5.6), then

$$S^2 > \left(\frac{c\Delta t}{\Delta \nu}\right)^2 \Rightarrow \left(\frac{\Delta \nu}{c\Delta t}\right)^2 > \frac{1}{S^2}. \tag{5.28}$$

And therefore (5.27) is larger than (5.21).

Comparing (5.27) **with** (5.24):

$$S^2 > \left(\frac{c\Delta t}{\Delta x}\right)^2 + \left(\frac{c\Delta t}{\Delta y}\right)^2 = \left(\frac{c\Delta t}{\Delta x}\right)^2 \cdot \left(1 + R_y^2\right). \tag{5.29}$$

Hence

$$\left(\frac{\Delta x}{c\Delta t}\right)^2 \cdot \frac{1}{1 + R_y^2} > \frac{1}{S^2}. \tag{5.30}$$

And thus (5.27) is larger than (5.24).

Comparing (5.27) **with** (5.25): Applying the same approach as in the previous case, it results that

$$\left(\frac{\Delta x}{c\Delta t}\right)^2 \cdot \frac{1}{1 + R_z^2} > \frac{1}{S^2}. \tag{5.31}$$

And thus (5.27) is larger than (5.25).

Comparing (5.27) **with** (5.26): Applying the same approach once again, it results that

$$\left(\frac{\Delta x}{c\Delta t}\right)^2 \cdot \frac{1}{R_y^2 + R_z^2} > \frac{1}{S^2}. \tag{5.32}$$

And thus (5.27) is larger than (5.26).

Since (5.27) is larger than any of the other extrema in 3D, it is thus clear that (5.27) is the maximum of the numerical group velocity. Then, the minimum of $\tilde{v}_g$ is determined by the minimum of equations (5.21), (5.24), (5.25) and (5.26). It then results that the minimum occurs only when one $\tilde{k}_\zeta$ is nonzero (i.e. along the axis with largest spatial step), Thus equation (5.21) is the minium. Furthermore, since these extrema are achieved for the same $\tilde{k}_\zeta$ as the numerical phase velocity, the propagation directions (ϕ, θ) of the extrema are the same as those presented in [32]. □

Note that, for cubic grid cells $\zeta_M = x, y, z$ and so the minimum is achieved $\forall\phi \in [0, \pi/2]$ when $\theta = 0$ and also for $\phi = 0$ and $\phi = \pi/2$ when $\theta = \pi/2$. On the other hand, the maximum occurs only for the propagation direction $(\phi, \theta) = (\pi/4, arctan(\sqrt{2}))$.

5.2.2 Extrema in 2D

As will be seen in the following, the solution in 2D results from a simplification of the 3D case. However, the resulting 2D formulas and their proofs are described next for the sake of completeness.

In 2D, the reduced equation for the numerical dispersion is

$$\frac{1}{(\Delta x)^2} sin^2\left(\frac{\tilde{k}_x \Delta x}{2}\right) + \frac{1}{(\Delta y)^2} sin^2\left(\frac{\tilde{k}_y \Delta y}{2}\right) = \frac{1}{(c\Delta t)^2} sin^2\left(\frac{\omega \Delta t}{2}\right) \tag{5.33}$$

and the Cartesian projections of the numerical wave vector and its module $\tilde{k}$ are

$$\tilde{k}_x = \tilde{k} cos(\phi) \tag{5.34}$$

$$\tilde{k}_y = \tilde{k} sin(\phi) \tag{5.35}$$

$$\tilde{k} = \sqrt{\tilde{k}_x^2 + \tilde{k}_y^2} \tag{5.36}$$

with ϕ the propagation angle with respect to the positive x direction. Similarly, the 2D Courant stability factor is

$$S = \sqrt{\left(\frac{c\Delta t}{\Delta x}\right)^2 + \left(\frac{c\Delta t}{\Delta y}\right)^2} \tag{5.37}$$

with $S \in [0, 1]$ for numerical stability. Further, the expression for the numerical group velocity in 2D is

$$\tilde{v}_g = A \cdot \left(\frac{\tilde{k}_x sin(\tilde{k}_x \Delta x)}{\tilde{k} \Delta x} + \frac{\tilde{k}_y sin(\tilde{k}_y \Delta y)}{\tilde{k} \Delta y}\right). \tag{5.38}$$

Theorem 2. *The maximum and minimum values of the numerical group velocity in standard 2D FDTD are*

$$\hat{v}_M = \frac{c}{cos(\frac{\omega \Delta t}{2})} \sqrt{1 - \frac{1}{S^2} sin^2\left(\frac{\omega \Delta t}{2}\right)} \tag{5.39}$$

$$\hat{v}_m = \min_{\zeta \in \{x,y\}} \left\{ \frac{c}{cos(\frac{\omega \Delta t}{2})} \sqrt{1 - \left(\frac{\Delta \zeta}{c \Delta t}\right)^2 sin^2\left(\frac{\omega \Delta t}{2}\right)} \right\} \tag{5.40}$$

which are achieved respectively for the propagation directions

$$\phi_{\hat{v}_M} = arctan\,(R_y) \tag{5.41}$$

$$\phi_{\hat{v}_m} = \begin{cases} 0, & if\, R_y \geq 1 \\ \frac{\pi}{2}, & if\, R_y \leq 1 \end{cases} \tag{5.42}$$

Proof. Similarly to the 3D case, the Lagrangian is

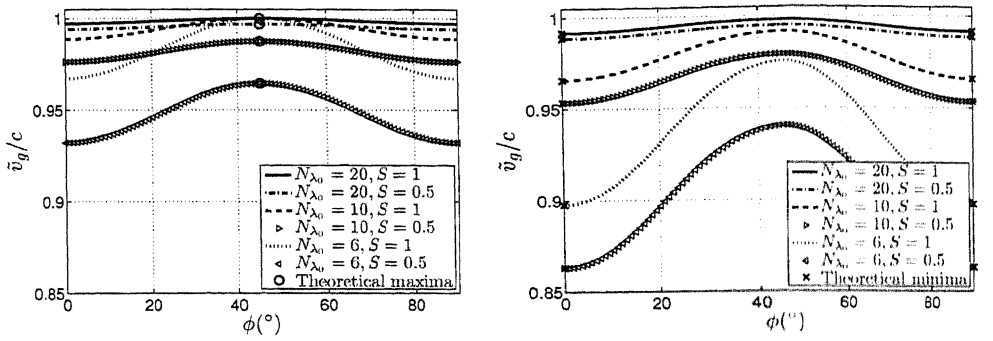
$$\begin{aligned} \mathcal{L}_g(\tilde{k}_x, \tilde{k}_y, \eta_g) = & \frac{\eta_g}{(c\Delta t)^2} sin^2\left(\frac{\omega \Delta t}{2}\right) + \\ & \left[\frac{A}{\Delta x} \frac{\tilde{k}_x sin(\tilde{k}_x \Delta x)}{\tilde{k}} - \frac{\eta_g}{(\Delta x)^2} sin^2\left(\frac{\tilde{k}_x \Delta x}{2}\right)\right] + \\ & \left[\frac{A}{\Delta y} \frac{\tilde{k}_y sin(\tilde{k}_y \Delta y)}{\tilde{k}} - \frac{\eta_g}{(\Delta y)^2} sin^2\left(\frac{\tilde{k}_y \Delta y}{2}\right)\right] \end{aligned} \tag{5.43}$$

The derivatives with respect to $\tilde{k}_x$ and $\tilde{k}_y$ in this case are those shown in equations (5.22) and (5.23). In the following and just as in the 3D case, let $\tilde{k}_\zeta$ and $\Delta\zeta$ with $\zeta \in \{x, y\}$ denote the different components of the wave vector and the spatial steps. Then, this system of equations has solutions under two different conditions:

Only one $\tilde{k}_\zeta$ is nonzero: In this case, let $\tilde{k}_\nu$ with $\nu \in \{x, y\}$ be the nonzero component of the wave vector, which can be directly obtained from the 2D dispersion relationship. Then, by plugging it into (5.38), this extremum of the numerical group velocity is the one shown in equation (5.21).

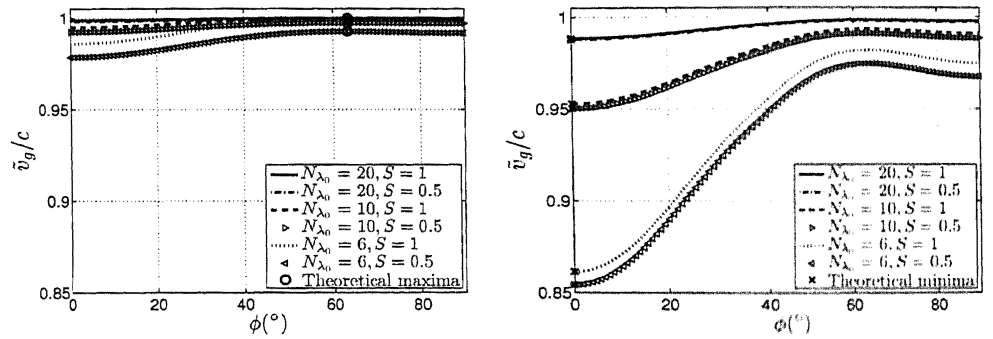
All $\tilde{k}_\zeta$ are nonzero: In this case, the system formed by equalling (5.22) and (5.23) to zero must be solved. This is done in Appendix A, where it is seen that the unique solution is $\Delta x \cdot \tilde{k}_x = \Delta y \cdot \tilde{k}_y$, thus leading to the value of the numerical group velocity shown in equation (5.27).

The results of (5.21) and (5.27) identify the extrema of $\tilde{v}_g$ in 2D. Then, in order to find out whether these are maximums or minimums, they must be compared. From (5.37), it is clear that equation (5.28) holds $\forall \nu \in \{x, y\}$, in view of which, (5.27) is the maximum of the numerical group velocity and (5.21) is the minimum.



(a) $\theta = arctan(\sqrt{1 + R_y^2}/R_z)$ (b) $\theta = 90°$

Figure 5.1: Normalised numerical group velocity in 3D FDTD with $N_{\lambda_0} = \lambda_0/\Delta x$, $R_y = 1$ and $R_z = 1$.



(a) $\theta = arctan(\sqrt{1 + R_y^2}/R_z)$ (b) $\theta = 90°$

Figure 5.2: Normalised numerical group velocity in 3D FDTD with $N_{\lambda_0} = \lambda_0/\Delta x$, $R_y = 2$ and $R_z = 3$.

Furthermore, these values are achieved for the same $\tilde{k}_\zeta$ as the numerical phase velocity (see [32] and [57]) and thus also for the same propagation angle ϕ. □

Also note that, for square grid cells, the minimum of the numerical group velocity is achieved at both angles 0 and $\pi/2$. On the other hand the maximum occurs only at $\phi = \pi/4$.

5.2.3 Numerical validation

The solution $\tilde{k}$ of the numerical dispersion relationship can be approximated numerically (e.g. using the Newton-Raphson method). Then, by substituting the

Table 5.1: Theoretical extrema of $\tilde{v}_g$ in 3D

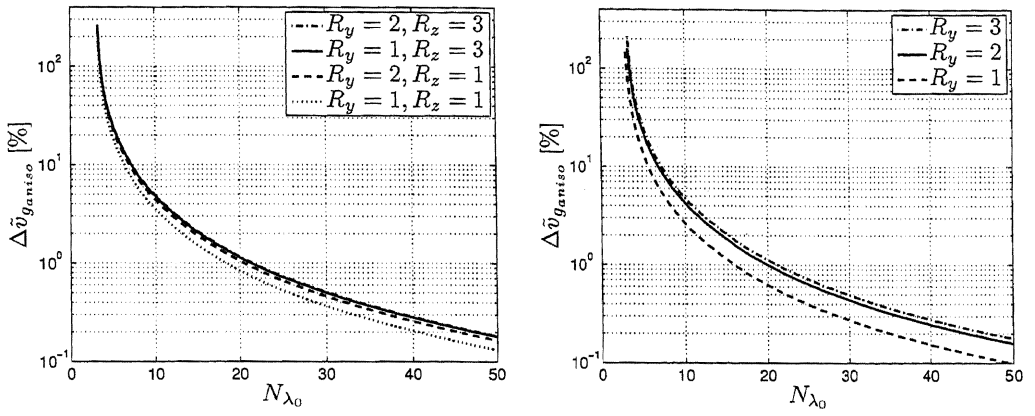
R_y	R_z	S	N_{λ_0}	$\hat{v}_m/c$	$\hat{v}_M/c$	$\Delta\tilde{v}_{g_{aniso}}$ [%]
1	1	1	20	0.9917	1.0000	0.8374
			10	0.9658	1.0000	3.5432
			6	0.8975	1.0000	11.4218
		0.5	20	0.9886	0.9969	0.8392
			10	0.9534	0.9875	3.5751
			6	0.8630	0.9646	11.7666
2	3	1	20	0.9885	1.0000	1.1670
			10	0.9528	1.0000	4.9484
			6	0.8614	1.0000	16.0857
		0.5	20	0.9878	0.9993	1.1676
			10	0.9502	0.9973	4.9579
			6	0.8543	0.9926	16.1869

obtained $\tilde{k}$ into (5.8) for 3D or into (5.38) for 2D, a good approximation to $\tilde{v}_g$ can be computed. Next, the theoretical extrema of $\tilde{v}_g$ obtained in the previous sections are validated. This is done by showing that the numerically computed $\tilde{v}_g$ is always bounded by these theoretical extrema.

In the 3D case, Figures 5.1 and 5.2 show numerically computed values of $\tilde{v}_g$ for different grid parameters, as well as the theoretical extrema according to Theorem 1. The angle resolution is $\Delta\phi = 0.09^o$ and $\tilde{k}$ is approximated using the Newton-Raphson method to an accuracy of 10^{-9}. In these examples $R_y \geq 1$, $R_z \geq 1$ and all magnitudes are expressed with respect to the wavelength corresponding to the carrier frequency f_0. Then, without loss of generality the grid resolution is defined as $N_{\lambda_0} = \lambda_0/\Delta x$. For convenience, the sine argument $\omega_0\Delta t/2$ is expressed next in terms of the FDTD grid parameters:

$$\frac{\omega_0\Delta t}{2} = \frac{2\pi c\Delta t}{2\lambda_0} = \frac{\pi}{N_{\lambda_0}}\frac{c\Delta t}{\Delta x} = \frac{\pi}{N_{\lambda_0}} \cdot \frac{S}{\sqrt{1+R_y^2+R_z^2}}. \tag{5.44}$$

This highlights that the extrema of $\tilde{v}_g$ and its anisotropy error in 3D depend mainly on R_y, R_z, S, and N_{λ_0}. Table 5.1 illustrates the theoretical values of the anisotropy error and the extrema. It can be seen that these match well the numerical results on Figures 5.1 and 5.2. The dependence of $\phi_{\hat{v}_M}$ with the cells aspect ratios is further



(a) 3D (b) 2D

Figure 5.3: Anisotropy error of the numerical group velocity when $S = 1$ and calculated with the theoretical formulas introduced in Theorems 1 and 2.

illustrated in Figure 5.2, where non-cubic cells are used. In addition, the anisotropy error in 3D and its dependence with the grid resolution is also illustrated in Figure 5.3(a). Finally and to give a global and clearer geometrical view of the behaviour of $\tilde{v}_g$, Figure 5.4 shows the numerically computed normalised group velocity values for a wider range of angular directions in the case of cubic cells with angle resolutions $\Delta\phi = 0.09^o$ and $\Delta\theta = 0.09^o$.

For the 2D case, Figure 5.5 shows numerically computed values of the normalised numerical group velocity $\tilde{v}_g/c$ in relationship with the propagation angle ϕ. For comparison purposes, the theoretical extrema of $\tilde{v}_g/c$ according to Theorem 2 are also plotted in this Figure. The angle resolution is $\Delta\phi = 0.09^o$ and $\tilde{k}$ is approximated using the Newton-Raphson method to an accuracy of 10^{-9}. Also, $R_y \geq 1$ and again without loss of generality the grid resolution is $N_{\lambda_0} = \lambda_0/\Delta x$. This way, the sine argument $\omega_0\Delta t/2$ can be expressed as

$$\frac{\omega_0\Delta t}{2} = \frac{\pi}{N_{\lambda_0}} \cdot \frac{S}{\sqrt{1+R_y^2}} \tag{5.45}$$

to highlight the main parameters that influence the anisotropy of $\tilde{v}_g$. In addition, Table 5.2 illustrates the theoretical values of the extrema and of the anisotropy error $\Delta\tilde{v}_{g_{aniso}}$. Then, simple visual inspection of Figure 5.5 and Table 5.2 reveals that the numerical results agree with the presented theory. This Figure also illustrates the

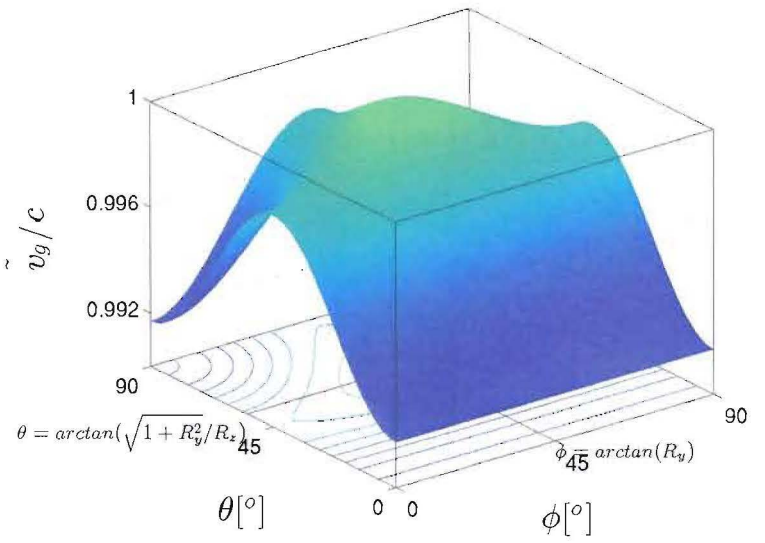


Figure 5.4: Normalised numerical group velocity in 3D FDTD with $N_{\lambda_0} = \lambda_0/\Delta x = 20$, $S = 1$ and $R_y = 1$ and $R_z = 1$.

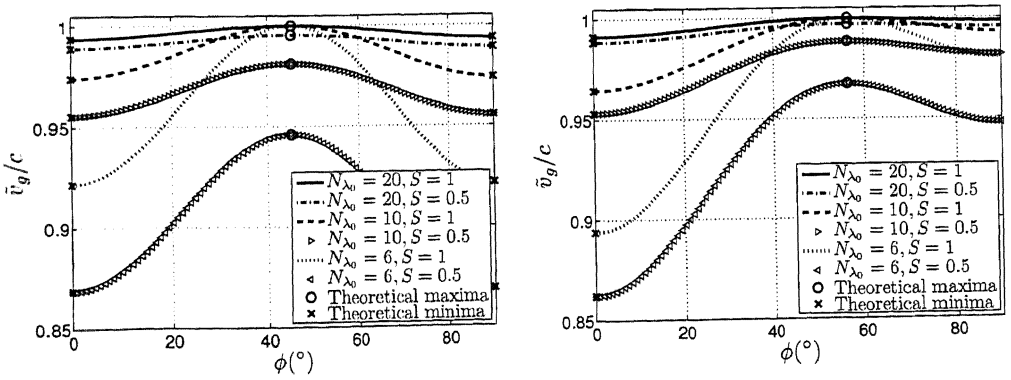
dependency of $\phi_{\hat{v}_M}$ with the cell aspect ratio R_y (see equation (5.41)) and, finally, the dependence of the anisotropy error with the grid resolution in 2D is illustrated in Figure 5.3(b).

5.3 Spread of numerical pulses

As seen in the previous section, the propagation of numerical wideband pulses is subject to various numerical artifacts. In particular and as a result of the anisotropy of the numerical group velocity, pulses propagate with different speeds along different propagation directions. Furthermore, pulses also spread in the time domain, thus reaching the receiving ends with a different amplitude shape than that of the source. This way, the pulse duration imposes a limit to the accuracy with which multipath components (MCs) in complex environments can be resolved. This phenomenon must hence be carefully considered when using pulsed sources to excite Yee grids for the purpose of channel impulse response prediction. In particular, for a fine resolution of the MCs, the duration of the received pulses must be much shorter than the expected delay spread. Hence, the quantification of this phenomenon is important for the configuration of the source and the grid properties. If the pulse

Table 5.2: Theoretical extrema of $\tilde{v}_g$ in 2D

R_y	S	N_λ	$\hat{v}_m/c$	$\hat{v}_M/c$	$\Delta\tilde{v}_{g_{aniso}}$ [%]
1	1	20	0.9938	1.0000	0.6278
		10	0.9742	1.0000	2.6529
		6	0.9216	1.0000	8.5069
	0.5	20	0.9891	0.9954	0.628
		10	0.9555	0.9812	2.653
		6	0.8686	0.9459	8.8994
1.5	1	20	0.9914	1.0000	0.8697
		10	0.9645	1.0000	3.6804
		6	0.8939	1.0000	11.8739
	0.5	20	0.9885	0.9971	0.8713
		10	0.9531	0.9885	3.7110
		6	0.8622	0.9674	12.2038



(a) $R_y = 1$ (b) $R_y = 1.5$

Figure 5.5: Normalised numerical group velocity in 2D FDTD with $N_{\lambda_0} = \lambda_0/\Delta x$.

spread could be predicted, wideband FDTD simulations could be properly set up to keep computational resources under control and the pulse spread bounded.

As a consequence of numerical dispersion, the numerical phase velocity of a given spectral component varies for different propagation directions. Such anisotropy has been already characterised for a wide range of finite-difference schemes, and several alternative methods have been proposed for its reduction [61]. However, the object of study here is different: this section focuses on the characterisation of the spread of numerical pulses due to the variation with frequency of the numerical phase velocity along a given propagation direction. For this, the phenomenon of Gaussian pulse spread due to numerical dispersion in Yee grids is first described and then quantified. Later, an analytical tool for the design of pulsed sources and Yee grids for wideband pulsed simulations is presented. Then and as will be shown, this helps to evaluate the accuracy with which multipath components from channel impulse response predictions can be resolved.

5.3.1 Spectral limit of wideband simulations

It can be seen from the dispersion equation of (5.1) that the relationship between $\tilde{k}$ and the angular frequency ω is non-linear. Hence, the numerical wave number $\tilde{k}$ of wideband signals propagating through the grid has a frequency spectrum. From [31], it is known that complex wave numbers give rise to exponential spatial decay of the propagated fields. This phenomenon is undesired and it can be reduced by controlling the spectral content of the source signal. It is thus important to realise that the source should not contain arbitrarily large frequency components in order to avoid complex $\tilde{k}$ values. The condition over the values of ω that lead to a complex wave number can be obtained from the extrema of $\tilde{k}$. From [57], these are

$$(\tilde{k})_m = \frac{2S}{c\Delta t} \cdot \arcsin\left(\frac{1}{S} sin\left(\frac{\omega\Delta t}{2}\right)\right) \tag{5.46}$$

$$(\tilde{k})_M = \max_{\zeta\in\{x,y,z\}} \left\{ \frac{2}{\Delta\zeta} \cdot \arcsin\left(\frac{\Delta\zeta}{c\Delta t} sin\left(\frac{\omega\Delta t}{2}\right)\right)\right\}. \tag{5.47}$$

In order for these not to be complex, the argument of the $arcsin(\cdot)$ functions must be less or equal to 1. Hence from (5.46)

$$f \leq \frac{1}{\Delta t} \cdot \frac{arcsin(S)}{\pi} \tag{5.48}$$

and from (5.47)

$$f \leq \frac{1}{\Delta t} \cdot \frac{arcsin\left(\frac{c\Delta t}{\Delta \nu}\right)}{\pi} \tag{5.49}$$

where f is the linear frequency and $\nu \in \{x, y, z\}$ is the dimension along which $\tilde{k}$ is maximum, which can be obtained by evaluating (5.47). Since (5.49) is more restrictive than (5.48), then (5.49) gives the maximum linear frequency component f_M that is not subject to spatial exponential decay in an FDTD simulation. Hence

$$f_M = \frac{1}{\Delta t} \cdot \frac{arcsin\left(\frac{c\Delta t}{\Delta \nu}\right)}{\pi}. \tag{5.50}$$

From this result, it can be seen that the spectral content of the source must be carefully considered when designing a wideband simulation. If this limit is not respected, the amplitude shape of the propagated wavepackets will be severely distorted due to the exponential spatial decay of over-the-limit spectral components.

5.3.2 The numerical phase delay

The speed of propagation of continuous wave (CW) signals of a given angular frequency ω is given by their phase velocity (see equation (4.3)) and this is true for both physical as well as for numerical waves. Since CW signals are nearly monochromatic, the calculation of $\tilde{v}_p(\omega)$ is then straightforward from the knowledge of ω and the numerical dispersion equation.

However, when bandpass wideband signals such as modulated pulses are used to excite the grid, the propagation speed of such pulses or wavepackets is not ruled by $\tilde{v}_p$ any more. In these cases, the numerical group velocity $\tilde{v}_g$, studied in the previous section, gives the speed with which pulses propagate through the grid. Further, the group velocity is a magnitude that makes only sense if the transmitted pulse maintains its shape approximately throughout the grid. Yet, it can be seen from the definition of phase velocity in equation (4.3) that each of the different frequency components that build the wideband pulse do not necessarily propagate with the same phase velocity. The result of this is an spread of the pulse in the time domain, which is akin to the phenomenon of *chromatic dispersion* in fibre optics. This effect is better understood by defining the numerical phase delay as

$$\tilde{T}_p(\omega) = \frac{d}{\tilde{v}_p(\omega)} = d \cdot \frac{\tilde{k}(\omega)}{\omega} \tag{5.51}$$

which gives the time that a spectral component with angular frequency ω takes to propagate a distance d through the grid. Then, for a pulse modulated at a carrier frequency ω_0 and with bandwidth $\Delta\omega$, the difference in the arriving time of the farmost frequency components is

$$\Delta\tilde{T}_p = d \cdot \left| \frac{1}{\tilde{v}_p(\omega_0 - \Delta\omega/2)} - \frac{1}{\tilde{v}_p(\omega_0 + \Delta\omega/2)} \right|. \tag{5.52}$$

If $\tilde{v}_p(\omega)$ was constant over all frequencies, or at least in the spectral region of width $\Delta\omega$ and centred on ω_0, then $\Delta\tilde{T}_p = 0$ and the pulse would retain its shape. However this is not necessarily true and the duration of the received pulse suffers an increase of $\Delta\tilde{T}_p$. Furthermore, it is seen from (5.51) that the numerical phase delay of a given spectral component achieves its extrema along the same direction as the extrema of $\tilde{k}$. A characterisation of such extrema is thus important and is hence developed in the following.

Theorem 3. *The extrema of the numerical phase delay $\tilde{T}_p(\omega)$ are monotonically increasing functions of ω.*

Proof. From (5.51), the definition of this theorem is equivalent to saying that the extrema of $\tilde{v}_p(\omega)$ are monotonically decreasing functions of ω, which is proven next.

In the following and without loss of generality let $R_y \geq 1$ and $R_z \geq 1$. Note that, due to the symmetry and periodicity of $\tilde{v}_p$ [32], other cell ratios are also considered by simply rotating the grid. The extrema of $\tilde{v}_p$ for an arbitrary spectral component are known and given by equations (4.5) and (4.6). Further, such expressions can be rewritten as

$$(\tilde{v}_p)_{max} = \frac{C_1 \cdot c}{\xi_1(a)} \tag{5.53}$$

and

$$(\tilde{v}_p)_{min} = \frac{C_2 \cdot c}{\xi_2(a)} \tag{5.54}$$

with

$$C_1 = \frac{1}{S} \geq 1, \tag{5.55}$$

$$C_2 = \frac{R_s}{S} > 1, \tag{5.56}$$

$$a = \frac{\omega\Delta t}{2} \tag{5.57}$$

and

$$\xi_i(a) = \frac{arcsin(C_i \cdot sin(a))}{a}. \tag{5.58}$$

It can be seen from (5.57) that a increases with the angular frequency ω. Moreover, $\xi_i(a)$ is a monotonically increasing function of a (see appendix C). Therefore, both $(\tilde{v}_p)_{max}$ and $(\tilde{v}_p)_{min}$ are monotonically decreasing functions of frequency. As a consequence, the extrema of the numerical phase delay are monotonically increasing functions of ω. □

Corollary 1. *Since $C_2 > 1$, then the maximum of the numerical phase delay is an strictly increasing function of frequency.*

Proof. If $C_2 > 1$ and following the process explained in Appendix C, $\xi_2(a)$ is an strictly increasing function of a. Thus, $(\tilde{v}_p)_{min}$ is an strictly decreasing function of frequency and hence, the maximum of $\tilde{T}_p$ is an strictly increasing function of frequency. □

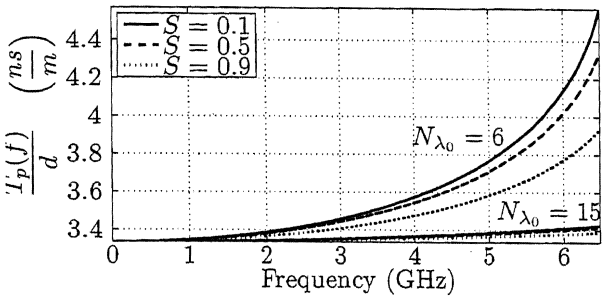
Figure 5.6 illustrates the behaviour of $\tilde{T}_p$ described by this theorem. Note that, in this Figure and in the following, the grid's spatial resolution is expressed with respect to λ_0, the physical wavelength corresponding to the carrier frequency f_0. Therefore,

$$N_{\lambda_0} = \frac{\lambda_0}{\Delta x}. \tag{5.59}$$

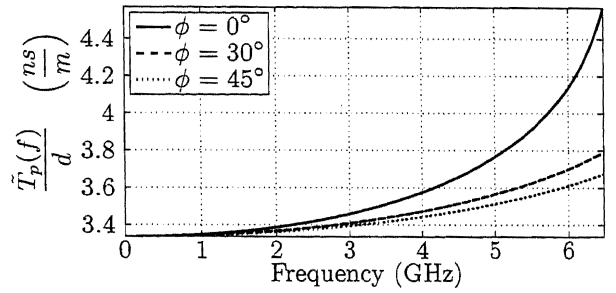
Since the focus of this chapter is on wideband simulations, the grid resolution N_λ could also be expressed with respect to an arbitrary spectral component of frequency f. Thus, N_λ is frequency-dependent and its value depends on the spectral component that it is referred to. However, the carrier frequency is used here for simplicity.

5.3.3 Linear approximation to the numerical phase delay

In order to obtain the numerical phase delay at a given frequency, the dispersion relationship (5.1) must be solved for $\tilde{k}$. Since (5.1) is a transcendental equation, an algebraic analysis of the dispersion suffered by a numerical pulse is not feasible. To make this task easier, a linear approximation to the phase delay is proposed next for an spectral band delimited by f_1 and f_2. This facilitates the analysis of the time spread of Gaussian pulses to be carried out in section 5.3.4. Then the accuracy of such an approximation will be evaluated in section 5.3.6. To highlight the applicability of this study to physical wideband signals, the linear frequency



(a) $\phi = 0$.



(b) $N_{\lambda_0} = 6$ and $S = 0.1$.

Figure 5.6: Frequency dependence of the numerical phase delay in 2D FDTD with square cells ($R_y = 1$) and $f_0 = 3.5$ GHz in a vacuum.

f instead of ω will be used in the following. The proposed approximation to the numerical phase delay of an spectral component with linear frequency f and after a propagated distance d from the source is thus

$$\hat{T}_p(f) = d \cdot (T_1 + m_T \cdot f) \quad f \in [f_1, f_2] \tag{5.60}$$

with $d \cdot T_1$ the numerical phase delay at $f = 0$ and $d \cdot m_T$ the slope of $\hat{T}_p(f)$. Thus, in order to fully characterise $\hat{T}_p(f)$, it is only necessary to determine T_1 and m_T.

Although $\tilde{T}_p(f)$ is undefined at $f = 0$, it can be proven (see Appendix D and Figure 5.6) that

$$\lim_{f \to 0} \tilde{T}_p(f) = \frac{d}{c}. \tag{5.61}$$

Thus $T_1 = 1/c$ can be used when $f_1 = 0$.

On the other hand, and as will be seen in section 5.3.4, m_T is the sole parameter responsible for the spread of a numerical pulse in an FDTD grid. Since $d \cdot m_T$ is

the slope of $\hat{T}_p(f)$, it is then reasonable, under the linear approximation of (5.60), to define $d \cdot m_T$ also as the slope of $\tilde{T}_p(f)$ at the carrier frequency. Thus

$$m_T \triangleq \frac{1}{d} \cdot \left. \frac{\partial \tilde{T}_p}{\partial f} \right|_{f=f_0} . \tag{5.62}$$

However, due to the well-known anisotropy of the phase velocity (see [31]), m_T is also subject to anisotropy and it thus varies with the propagation direction. Since m_T is the main responsible for the spread of numerical pulses, it is thus important to characterise its bounds.

From the definition of numerical phase delay in (5.51) and from (5.62), m_T can be written as

$$m_T = \frac{1}{\omega \cdot \tilde{v}_g} \cdot \left(1 - \frac{\tilde{v}_g}{\tilde{v}_p}\right) . \tag{5.63}$$

In [57], it was shown that $\tilde{v}_g/\tilde{v}_p$ is maximum for the direction $(\phi_{\tilde{k}_{min}}, \theta_{\tilde{k}_{min}})$, and that it is minimum for $(\phi_{\tilde{k}_{max}}, \theta_{\tilde{k}_{max}})$. Furthermore, it was proven in section 5.2 that $\tilde{v}_g$ is maximum for $(\phi_{\tilde{k}_{min}}, \theta_{\tilde{k}_{min}})$, and that it is minimum for $(\phi_{\tilde{k}_{max}}, \theta_{\tilde{k}_{max}})$. Thus, it can be inferred from (5.63) that m_T achieves its extrema on the same direction as the numerical wave number $\tilde{k}$.

The implications of this result is that the pulse spread is maximum along $(\phi_{\tilde{k}_{max}}, \theta_{\tilde{k}_{max}})$. Thus, the analysis of the source design can be simplified by considering only such direction.

5.3.4 Time-spread of a Gaussian pulse under linear phase delay

If the numerical phase delay were linearly dependent on frequency, the analytical study of the temporal spread of pulses could be simplified. To study this, let us first define the Gaussian source in the time-frequency domain as

$$p_{tx}(t, f) = A \cdot e^{-\left[\left(\frac{t-t_0}{t_w}\right)^2 + (t_w \cdot f)^2\right]} \tag{5.64}$$

where A is the amplitude of the pulse, t_0 the instant of maximum amplitude and t_w the pulse duration measured at a power[1] drop of $-20 \cdot log_{10}(e^{-(1/2)^2}) \approx 2.17$

[1]The power of a pulse $p(t)$ in logarithmic units is $P(t) = 20 \cdot log_{10}(p(t))$.

dB below the maximum. Then, the spectral components of the received pulse p_{rx} undergo frequency-dependent phase delay. Hence

$$p_{rx}(t,f) = p_{tx}(t - \hat{T}_p(f), f) \tag{5.65}$$

with $\hat{T}_p(f)$ the linear approximation of (5.60) to the numerical phase delay. In order to characterise the effect of this delay on the pulse in the time-domain, $p_{rx}(t,f)$ is projected onto the time dimension by integrating over the frequency. Thus

$$p_{rx}(t) = \int_{-\infty}^{\infty} p_{rx}(t,f)df. \tag{5.66}$$

This yields

$$p_{rx}(t) = \frac{A}{\sqrt{\pi \cdot \left[t_w^2 + (\frac{m_T \cdot d}{t_w})^2\right]}} \cdot e^{-\left(\frac{t - t_0 - d \cdot T_1}{\sqrt{t_w^2 + \left(\frac{m_T \cdot d}{t_w}\right)^2}}\right)^2}. \tag{5.67}$$

The duration of the received pulse in this expression is given by the denominator in the exponent of e. Therefore, for a source Gaussian pulse of duration t_w, the duration t'_w of the received numerical pulse at a distance d is

$$t'_w = \sqrt{t_w^2 + \left(\frac{m_T \cdot d}{t_w}\right)^2}. \tag{5.68}$$

This expression highlights that the slope $m_T \cdot d$ of the numerical phase delay is the main parameter that affects the time spread of numerical pulses. For illustration, the relationship between the source duration and the received pulse duration is presented in Figure 5.7. From this, it may seem counterintuitive that two different sources (e.g. one with duration t_{w1} and another with duration t_{w2}) produce a pulse of the same duration t'_w at a distance d. However, this can be explained in terms of their frequency contents: Sources with large bandwidths (i.e. low t_w) suffer more dispersion than sources with lower bandwidths. That is why t'_w tends to infinity as t_w tends to zero. On the other hand, as the bandwidth of the source diminishes (i.e. larger t_w), $\Delta\tilde{T}_p$ also decreases. Ideally, as the bandwidth tends to zero (t_w tends to infinity), all spectral components within such a narrow band arrive with the same delay and the pulse does not spread. Hence, for long pulses (large t_w), it results that $t'_w \approx t_w$.

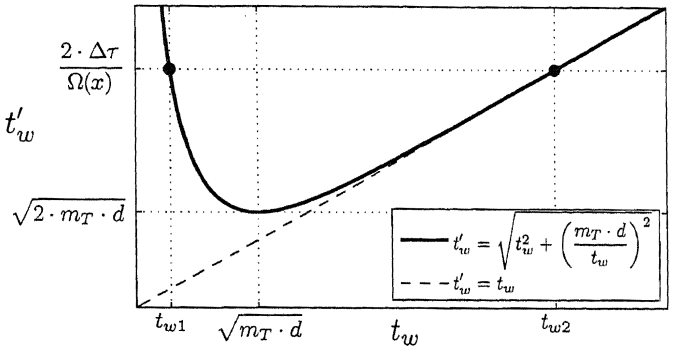


Figure 5.7: Relationship between the duration t'_w of a received pulse and the duration t_w of the source. The values of t_w that bound the interval where $t'_w < (t'_w)_{max}$ (see equation (5.73)) are indicated with dots. The asymptote $t'_w = t_w$ is shown for reference.

5.3.5 Source design for controlling the spread of numerical pulses

The objective here is to design a Gaussian source, such that the temporal spread of the received pulses is bounded. This way and given a set of time and power constraints, adjacent pulses do not overlap in time and can thus be resolved. To do this, the shape parameter t_w of the source that matches these conditions must be carefully chosen. Let us then define a modulated Gaussian pulse as

$$p_{tx}(t) = A \cdot e^{-\left(\frac{t-t_0}{t_w}\right)^2} \cdot sin\left(2\pi f_0 t\right) \tag{5.69}$$

where f_0 is the carrier frequency and A is the amplitude. Just as before, t_w is the duration of the pulse measured 2.17 dB below the power maximum. Figure 5.8 shows an example of a Gaussian pulse in which the duration t_{w_x}, measured at an amplitude decay factor of x with respect to the maximum, is also indicated. For pulse design considerations, the relationship between t_{w_x} and t_w is

$$\frac{t_{w_x}}{t_w} = \Omega(x) \tag{5.70}$$

with

$$\Omega(x) = 2 \cdot \sqrt{-\ln\left(\frac{1}{x}\right)}. \tag{5.71}$$

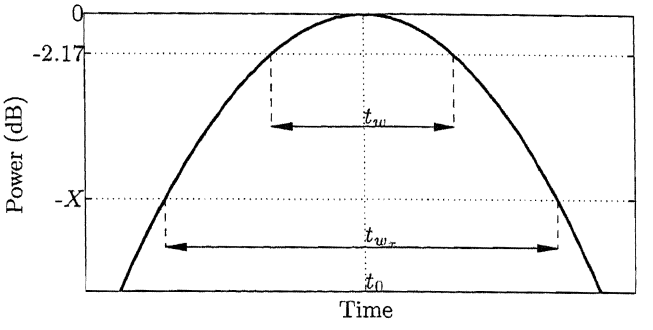


Figure 5.8: Non-modulated Gaussian pulse ($f_0 = 0$) in the time domain.

When a Gaussian pulse in an FDTD simulation is received far from the source, equation (5.68) showed that the duration of the pulse is larger than that of the originally transmitted pulse. Hence, a numerical Gaussian pulse received after travelling over a straight line of length d from the source reaches an amplitude x times lower than the maximum only after a period of duration $t'_{w_x}/2$ after the peak. Therefore, replacing t_w by (5.68) in (5.70) yields

$$t'_{w_x} = \Omega(x) \cdot \sqrt{t_w^2 + \left(\frac{m_T \cdot d}{t_w}\right)^2} \tag{5.72}$$

which is the duration of the received numerical pulse measured $X = 20 \cdot log(x)$ dB below the power maximum (see Figure 5.8).

The target source must thus give rise to the maximum desired temporal duration t'_{w_x} of the received pulse after a travelled distance d and for a custom attenuation x. The maximum allowed t'_w is hence given in terms of the required minimum time separation $\Delta\tau$ (see Figure 5.9) between pulses and the amplitude decay factor x. Thus,

$$(t'_w)_{max} = \frac{(t'_{w_x})_{max}}{\Omega(x)} = \frac{2 \cdot \Delta\tau}{\Omega(x)}. \tag{5.73}$$

Hence, in order to resolve adjacent pulses under these conditions, the source parameter t_w must comply with the following inequation:

$$t'_w \leq (t'_w)_{max} \tag{5.74}$$

or equivalently, using (5.68) and (5.73):

$$\sqrt{t_w^2 + \left(\frac{m_T \cdot d}{t_w}\right)^2} \leq \frac{2 \cdot \Delta\tau}{\Omega(x)}. \tag{5.75}$$

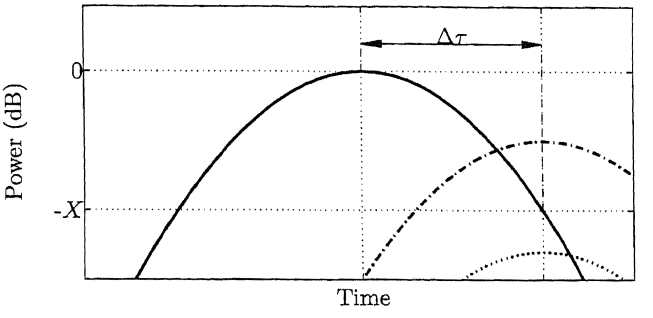


Figure 5.9: The solid line represents the strongest received pulse. The dash-dotted line is a resolvable pulse within a power range of X dB from the maximum. The dotted line represents a non-resolvable pulse.

Note that in order for the time separation $\Delta\tau$ between resolvable pulses to be valid in all the scenario with the source designed according to (5.75), the expected maximum travel distance d of a pulse must be given. This method is thus capable of predicting the maximum spread of pulses after propagating over a distance smaller than d.

There is a full range of t_w values that meet inequation (5.75). To find them, the corresponding equation is defined by replacing $\leq$ with $=$. Then, a polynomial equation of degree four results. There are thus four different solutions. Neglecting the two negative solutions, there are still two positive values of t_w that satisfy it, thus defining the range of solutions of (5.75). These are

$$t_{w1} = \sqrt{\frac{B - \sqrt{B^2 - (2 \cdot m_T \cdot d)^2}}{2}} \tag{5.76}$$

$$t_{w2} = \sqrt{\frac{B + \sqrt{B^2 - (2 \cdot m_T \cdot d)^2}}{2}} \tag{5.77}$$

with

$$B = (t'_w)^2_{max}. \tag{5.78}$$

However, these only exist if the conditions over $(t'_w)_{max}$ are not too strong. This is clearly illustrated on Figure 5.7, which shows that m_T defines the minimum duration t'_w that can be achieved by a pulse received at a distance d. Hence, solutions to (5.75) exist if and only if

$$\frac{2 \cdot \Delta\tau}{\Omega(x)} \geq \sqrt{2 \cdot m_T \cdot d}. \tag{5.79}$$

Therefore, any source with $t_w \in (t_{w1}, t_{w2})$ is guaranteed to meet the time and power resolution requirements of (5.74), as long as (5.79) is respected. If the minimum of t'_w as shown in Figure 5.7 is not enough for a given application, m_T as defined in (5.62) can be tuned by adjusting the grid properties. This usually involves augmenting the grid resolution and thus the computational cost.

Finally, note that the objective in this section is to obtain the source duration t_w for a given time resolution of pulses. The procedure for the design of Gaussian sources that guarantee maximum levels of pulse spread is then summarised in the following steps:

1. Specification of time resolution requirements: Two pulses are resolvable if the peak of the faintest one is above the level of the strongest one. This is illustrated in Figure 5.9, in which $\Delta\tau$ is the minimum time delay necessary to resolve pulses with amplitudes that differ by a factor of x.

2. Estimation of the maximum distance d to be travelled by the pulses: Since multipath is to be considered, this can be done by calculating the necessary distance for an attenuation factor x in the background medium.

3. If f_0 and S are given:

   (a) Calculate m_T with (5.62).

   (b) If (5.79) is true:

      i. Calculate the bounds of the required source duration t_w with (5.76) and (5.77).

      ii. Choose $t_w \in [t_{w1}, t_{w2}]$.

   (c) If (5.79) is false, there is no source satisfying the conditions.

4. If f_0 and/or S are unknown:

   (a) Calculate the maximum required m_T with (5.79).

   (b) Adjust f_0 and/or S with (5.62) so that m_T is below or equal to the limit of (5.79).

   (c) Recalculate m_T with the chosen f_0 and S.

(d) Calculate the bounds of the required source duration t_w with (5.76) and (5.77).

(e) Choose $t_w \in [t_{w1}, t_{w2}]$.

5.3.6 Numerical validation

The application of interest for the theory just presented is the modelling of radio channels in electrically large areas. More specifically, in femtocell scenarios (indoors-to-outdoors). At the frequencies used by current mobile telecommunication systems (e.g. $0.9 - 3.5$ GHz), this is feasible using 2D simulators such as the one described in section 3.3. Hence, the purpose of this section is to evaluate the degree of accuracy with which the duration t'_w of a received pulse can be predicted in 2D using expression (5.68). Note that this formula results as a consequence of applying the linear approximation (5.60) to the numerical phase delay. Therefore, the predicted t'_w is subject to an error that depends on the validity of such a linear approximation. From Figure 5.6(a), it can be seen that $\tilde{T}_p$ presents a more linear behaviour at higher N_{λ_0} and S values. It is thus expected that the accuracy with which t'_w can be predicted will also be higher for larger grid resolutions.

Furthermore, it has been shown in section 5.3.3 that the time spread of a pulse is maximum along the direction on which $\tilde{k}$ is maximum. Such direction is $\phi_{(\tilde{k})_M}$ in 2D or $(\phi_{(\tilde{k})_M}, \theta_{(\tilde{k})_M})$ in 3D. The assessment performed here is thus restricted to this direction, as the pulse duration along any other direction is always lower. Therefore, from (5.62) and for the case of square ($R_y = 1$) and/or cubic ($R_y = 1$ and $R_z = 1$) grid cells, ($\phi_{(\tilde{k})_M} = 0\,°, \theta_{(\tilde{k})_M} = 0\,°$) (see [32]) and thus the largest slope m_T of the linear approximation to the numerical phase delay is

$$(m_T)_M = \frac{1}{c \cdot f_0} \cdot \left[\frac{cos\left(\frac{\omega_0 \Delta t}{2}\right)}{\sqrt{1-\left(\frac{R_s}{S}\right)^2 \cdot sin^2\left(\frac{\omega_0 \Delta t}{2}\right)}} - \frac{N_{\lambda_0}}{\pi} \cdot arcsin\left(\frac{R_s}{S} \cdot sin\left(\frac{\omega_0 \Delta t}{2}\right)\right) \right]. \quad (5.80)$$

In the following, FDTD simulations are performed in order to compare the theoretical pulse duration $t'_w|_{th}$ as predicted by (5.68) with (5.80) against the pulse duration $t'_w|_{sim}$ measured from the simulated numerical pulses. Once $t'_w|_{sim}$ is measured and $t'_w|_{th}$ calculated, the relative error in the theoretical prediction can be

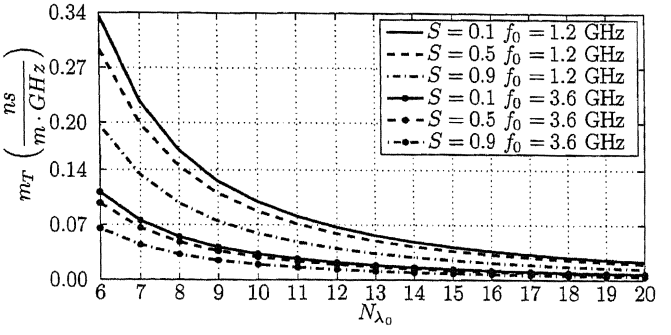


Figure 5.10: Relationship between the slope m_T of the linear approximation $\hat{T}_p$ as defined in (5.80) with other grid parameters.

defined as

$$E = \frac{\left|t'_w|_{sim} - t'_w|_{th}\right|}{t'_w|_{th}}. \tag{5.81}$$

Hence this quantifies the precission with which (5.68) and (5.80) can predict the pulse spread in 2D FDTD simulations with square grid cells.

5.3.6.1 Measuring the duration of a numerical pulse

In order to compare the computed pulse duration against the theoretical predictions, it is necessary to measure $t'_w|_{sim}$ from the FDTD simulations. This is however not a trivial task and different methods can be applied. For instance, the extraction of power delay profiles (PDPs) from simulation results has been used [59], as well as non-coherent demodulation techniques [22].

Since the objective here is to design bandpass sources, the duration of a numerical pulse $\tilde{p}$ is measured from its amplitude envelope, which is equivalent to the square root of its power delay profile. The amplitude envelope $\tilde{p}_{env}$ of a real bandpass signal $\tilde{p}$ is that of the complex analytic signal (also called the pre-envelope of $\tilde{p}$), which is defined as

$$\tilde{p}_a = \tilde{p} + j \cdot \tilde{p}_H \tag{5.82}$$

where j is the imaginary unit and $\tilde{p}_H$ is the Hilbert transform of $\tilde{p}$ [62]. Thus

$$\tilde{p}_{env} = \|\tilde{p}_a\| = \sqrt{\tilde{p}_a \cdot \tilde{p}_a^*} = \sqrt{\tilde{p}^2 + \tilde{p}_H^2}. \tag{5.83}$$

Then, the process for the measurement of the duration $t'_w|_{sim}$ of a numerical pulse $\tilde{p}$ received at a given location in the grid is:

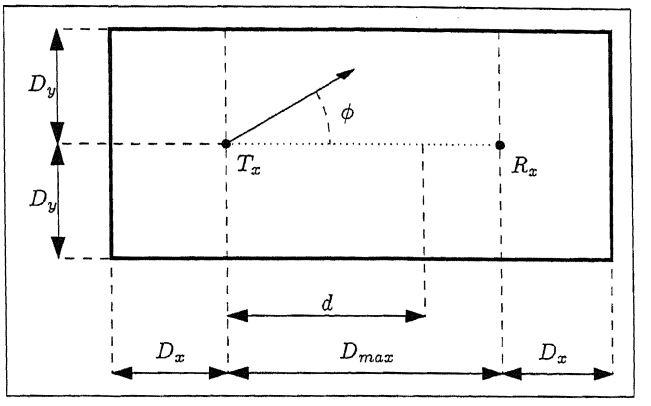


Figure 5.11: Scenario configuration for numerical validation.

1. Computation of the Discrete Hilbert transform $\tilde{p}_H$ of $\tilde{p}$.
2. Extraction of the amplitude envelope $\tilde{p}_{env}$ with (5.83).
3. Measurement of the peak amplitude of $\tilde{p}_{env}$.
4. Measurement of $t'_w|_{sim}$ as the time difference between the samples at which the amplitude of $\tilde{p}_{env}$ is $e^{-(1/2)^2}$ times lower than the peak (equivalent to measuring t_w in Figure 5.8).

5.3.6.2 Simulation settings

To assess the presented theory, a series of scenarios have been designed to evaluate the corresponding error as defined in (5.81). The grid is an staggered 2D Yee lattice and the scheme is the standard FDTD in TM_Z mode. The cells are square ($R_y = 1$) and there is a soft source exciting the grid through the generation of a numerical electric current J_s, which is a modulated Gaussian pulse

$$J_s(t) = e^{-\left(\frac{t-t_0}{t_w}\right)^2} \cdot sin(2\pi f_0 t). \tag{5.84}$$

The simulation scenario is illustrated in Figure 5.11, where the transmitter is indicated by T_x and the furthest measurement location at a distance D_{max} from the source is marked by R_x. Since the pulses travel the grid at the numerical group velocity $\tilde{v}_g$ which is known (see section 5.2), the scenario dimensions (D_x, D_y) for

a given D_{max} can be adjusted to avoid the time-domain overlapping of pulses at receiving points at any distance $d < D_{max}$ along the line $\phi = 0°$.

Overlapping is avoided by guaranteeing a minimum time separation $\Delta\tau$ on reception between the pulse that travels along $\phi = 0°$ and the reflections on the boundaries. Pulses reflected on the left and right borders travel always along $\phi = 0°$, while pulses reflected on the upper/lower boundaries travel with constant group velocity along an arbitrary direction ϕ ($\tilde{v}_g(\phi)$ has a period of 90°). In the case of square grid cells, the group velocity is minimum on $\phi = 0°$. Using then basic geometry for the case when the pulse along $\phi = 0°$ arrives first, the scenario size must satisfy the following conditions:

$$D_x > \frac{\Delta\tau}{2} \tag{5.85}$$

$$D_y > \frac{1}{2} \cdot \sqrt{\left[\tilde{v}_g(\phi) \cdot \left(\frac{d}{(\tilde{v}_g)_{min}} + \Delta\tau\right)\right]^2 - d^2}. \tag{5.86}$$

If the maximum distance D_{max} to be measured is known, then a conservative limit on (5.86) can be obtained by substituting d by D_{max} and $\tilde{v}_g(\phi)$ by $\hat{v}_M$. Then, the grid size (N_x, N_y) in number of cells can be calculated from the knowledge of the spatial step. Finally, the minimum number N_t of necessary time steps for the arrival of the pulse at the furthest measurement location can be estimated by

$$N_t \geq \frac{1}{\Delta t} \cdot \left(t_0 + \frac{t'_{w_x}}{2} + \frac{D_{max}}{(\tilde{v}_g)_{min}}\right) \tag{5.87}$$

where Δt is the simulation time step and t'_{w_x} is the expected received pulse duration that depends on the amplitude drop to be measured. Since the objective in this case is to measure t'_w for comparison purposes, then $t'_{w_x} = t'_w$ was used.

The numerical pulses are then measured along a straight line on the propagation direction $\phi = 0°$ and the measurement locations are spaced Δd. As explained in chapter 4, the FDTD simulator is implemented on a graphics processing unit with memory limitations. Therefore, the largest simulated grid was $(N_x, N_y) = (2368, 1968)$ and the furthest measurement location $D_{max} = 19$ m. Table 5.3 collects the range of other simulation parameters that have been tested.

Table 5.3: Simulation settings for t'_w prediction

Parameter	Tested values
f_0	1.2 GHz
t_0	6 ns
S	$0.1, 0.5, 0.9$
N_{λ_0}	$6, ..., 15$
t_w	1.5, 2, 2.5 ns
d	[5, 19] m
D_{max}	19 m
Δd	5 cm

5.3.6.3 Simulation results

As explained above, the pulse duration measurement procedure of section 5.3.6.1 is not perfect. Hence, due to inaccuracies in the measurement of t'_w from the received numerical pulses, the error (5.81) is subject to random variations among different locations. For this reason, Figure 5.12 shows the percentage errors as a standard box-and-whisker diagram. This way, each box represents the distribution of the error along all measurement points for a given simulation settings.

One immediate conclusion of these results is that the prediction of t'_w obtained with (5.68) is more accurate for higher grid resolutions. This was expected, as $\tilde{T}_p$ tends to linearity for increasing N_{λ_0}. Furthermore, increasing Courant numbers also reduce the prediction error. However, it can be observed from Figure 5.12 that E is weakly dependent on S.

Finally, it is also seen that the prediction is more accurate for longer pulses (sources with larger t_w). Since longer pulses have narrower bandwidths, this effect is also explained in terms of the frequency content of the source. $\tilde{T}_p(f)$ is not a linear function of f. However, the error of the slope of the linear approximation of $\tilde{T}_p(f)$ is zero at $f = f_0$. This is because m_T is defined as the derivative of $\tilde{T}_p(f)$ at $f = f_0$ (see (5.62)). Furthermore, this error increases as frequencies depart from f_0. Hence, narrower bandwidths reduce the error in the estimation of m_T and thus, also of the pulse spread.

Equation (5.68) predicts the existence of a source that causes a pulse received at

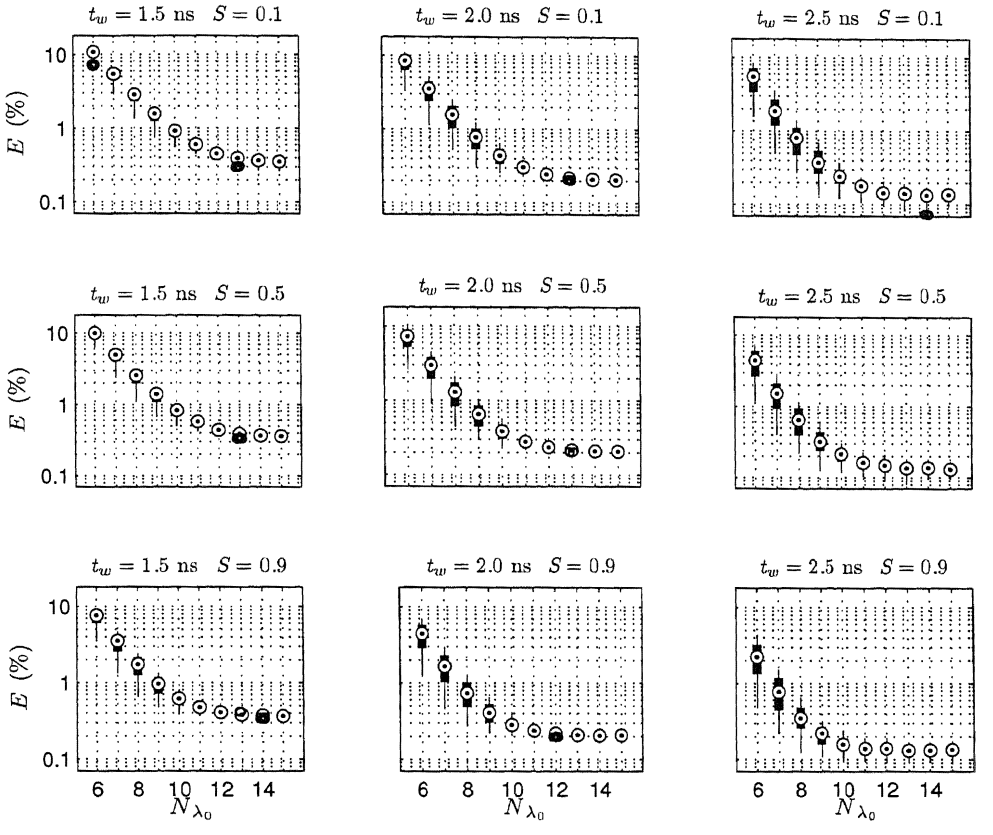


Figure 5.12: Box-and-whisker diagram representing the percentage error E of the t'_w predictions under different simulation conditions. The box length is the interquartile range and the median is indicated by a black dot in a white circle. The whiskers maximum length is 1.5 times the interquartile range and any E value beyond the end of the whiskers is marked by a circle (outliers).

a distance d to have a minimum duration $\sqrt{2 \cdot m_T \cdot d}$. By using sources of different durations t_w and measuring t'_w at different distances, this behaviour has been observed and thus confirmed in the numerical simulations. This is illustrated in Figure 5.13 where the higher accuracy of the prediction at larger N_{λ_0} values is also evident.

5.3.7 Pulse propagation in complex scenarios

For simplicity, the concepts and analysis presented in this section are restricted to pulses propagating in one medium only. Hence, these results can be used to control

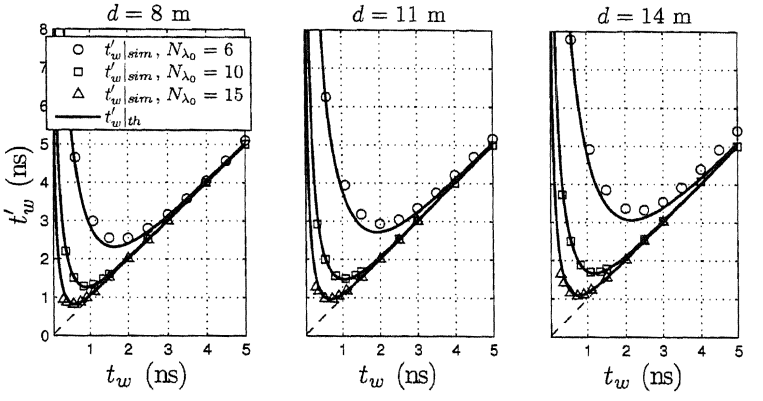


Figure 5.13: Simulated and predicted pulse durations at different distances d from the source. The simulation parameters are $S = 0.1$ and $f_0 = 1.2$ GHz.

the spread of numerical pulses propagating in the background material. However, if the simulation environment contains several materials, additional phenomena need to be considered:

- **Reflections:** When a pulse propagating in the background medium encounters an obstacle, a fraction of its energy is reflected back into the same medium. Hence, the time spread suffered by pulses that undergo multipath reflections can still be controlled by considering the spread along the direction with largest m_T.

- **Refractions:** After encountering an obstacle, not only is a fraction of the energy reflected but some is also transmitted to the new medium. If the refraction index of the obstacle is different from that of the first medium, the numerical phase velocity changes and so does the numerical phase delay. Materials with larger refraction indices are thus subject to larger numerical phase delays. Furthermore, $\tilde{T}_p$ is proportional to the refraction index and thus, so is m_T. Hence, the spread of numerical pulses is larger in materials with larger refraction indices. This can also be quantified with the model presented here as long as the material properties are known.

5.4 Conclusions

This chapter has characterised, for the first time, the two main sources of error in FDTD simulations of wideband pulsed sources. Because of these studies, Yee grids can now be configured such that a minimum level of accuracy is guaranteed when the computational resources are limited.

In section 5.2, analytical expressions for the extrema of the numerical group velocity in 3D and 2D Yee grids of arbitrary cell sizes were obtained. Further, it was shown that the extrema of $\tilde{v}_g$ occur along the same directions as the extrema of $\tilde{v}_p$. Based on this, the term *special direction* used by Zhao in [32] acquires higher significance, as these directions present indeed *special* properties for both wideband and CW signals. In addition, these results have fully characterised the anisotropy error of the numerical group velocity. Then, equations (5.10) and (5.39) also revealed that $\tilde{v}_g$ can be faster than c (faster-than-light or superluminal numerical pulse propagation) when $S > 1$. However, such solutions are known to be numerically unstable [31].

Further, it was also revealed that the anisotropy error depends not only on the grid properties but also on the speed of propagation c in the medium being modelled. Thus, if simulations containing different media are considered, $\Delta\tilde{v}_{g_{aniso}}$ will be larger within those materials that have a lower propagation speed (i.e. a higher refractive index). Nevertheless, the general formulation of the derived equations can be applied to predict the anisotropy error in all media involved in the simulation, given that the values of c for each material are known. The grid parameters can then be adjusted to guarantee a maximum anisotropy error in the medium with the lowest propagation speed.

Then, in section 5.3, a theory on the anisotropy of the spectral variation of the numerical phase delay was presented. This characterised several aspects related to the numerical dispersion of wideband pulses when propagating through Yee grids. Specifically, formulas for the prediction of the pulse spread and based on a linear approximation to $\tilde{T}_p$ were introduced and validated through simulations. The propagation directions along which the pulse spread reaches its maximum were also identified. This result simplifies the source design process, because only those directions along which the spread is maximum need to be evaluated for spread control.

Then, the necessary steps that must be followed to keep the pulse spread bounded and the resolution of pulses high have been properly described and summarised.

Finally, it was shown that the time-domain spread of numerical pulses in FDTD is due to the variation of the numerical phase delay with frequency. In addition, an upper boundary to the spectral content of wideband sources was also introduced. If this limit is not respected, the amplitude shape of the numerical pulses will be severely distorted due to exponential spatial decay of high frequency spectral components.

5.4.1 Contributions

Finally and to summarise, the main contributions of this chapter made by the author to the scientific community are summated next and illustrated with numerical examples:

1. **Analytical formulae to predict the propagation directions and values of the extrema of the numerical group velocity:** Thanks to this, the anisotropy error of $\tilde{v}_g$ in wideband simulations in non-lossy media is now fully characterised. Furthermore and with view to channel impulse response prediction, a bound to the relative error with which the time of arrival of multipath components is predicted using FDTD simulations of pulsed sources can now be established.

   To illustrate this, Figure 5.3 showed that 2D wideband simulations with square grid cells are subject to an anisotropy of the numerical group velocity close to 11% when $N_{\lambda_0} = 6$ and $S = 1$. This anisotropy error also measures the uncertainty in the predicted ToA of MCs in a multipath scenario. Thus, let such a FDTD simulation predict that a multipath component reaches the receiver after 100 ns of propagation. Then, it can only be guaranteed that the corresponding physical multipath component will reach the receiver at some moment between $100/1.11 \approx 90$ ns or $100 * 1.11 \approx 111$ ns of propagation time.

2. **Analytical formulae to approximate the time spread of numerical pulses:** These can be used to design wideband simulations in which the duration of the received numerical pulses is bounded. This means that the designer of the simulation can choose a duration for the received pulse according to its

own needs. Hence, the resolution of multipath components that are separated by a minimum time of arrival is guaranteed and wideband parameters of radio channels can be predicted with the desired level of accuracy.

The applicability of this result is best illustrated in Figure 5.13. This shows that every grid defines a pulsed source that results in the minimum pulse duration on the receiver end. For instance, let the furthest receiver be located 14 meters from the transmitter and the grid be configured as $N_{\lambda_0} = 10$ and $S = 0.1$. It is then seen from Figure 5.13 that it does not make sense to use a source pulse shorter than 1 ns, because the received pulse would be far longer than the one achieved with a 1 ns source pulse. The received pulse here lasts approximately 1.8 ns. If a shorter pulse is needed (e.g. in femtosecond laser applications), there is no other choice (with this FDTD scheme) than increasing the grid resolution.

The practical application of the theoretical results introduced here is illustrated in the next chapter for the characterisation of the indoor-to-outdoor radio channel.

6

Wideband channel modelling of femtocell scenarios with the finite-difference time-domain method

This chapter illustrates the applicability of the FDTD method to the prediction of wideband parameters in radio channels. The theory presented in sections 5.2 and 5.3 of the previous chapter is used next to guarantee the validity of the predictions in terms of the time of arrival (ToA) and the resolvability of multipath components (MCs). For this, the grid will be excited with a Gaussian pulse designed to achieve a specific degree of accuracy in the resolution of MCs. Then, the channel impulse response (CIR) is recorded on several locations under different setups and at a variety of carrier frequencies: 2.36 GHz in an indoor scenario, and 0.9, 2.5 and 3.5 GHz in two indoor-to-outdoor scenarios. The size of the largest scenario is $(D_x, D_y) = (77, 47)$ meters (see Figures 6.2 and 6.3), needing this to be considered when sampling the spatial grid.

6.1 Simulation Setup

As previously seen and due to numerical dispersion, FDTD simulations of wideband pulses are subject to two fundamental distortions:

1. Anisotropy of the numerical group velocity.

2. Temporal spread of pulses.

Hence, the settings of the simulations to be performed in this section are based on the previously presented theory with the objective of monitoring and controlling the errors arising from these phenomena. These settings are explained in the following.

6.1.1 Control over the numerical group velocity

In order to minimise $\Delta\tilde{v}_{g_{aniso}}$, a proper spatial sampling N_{λ_0} has to be chosen. For this, it was seen in section 5.2 that $\Delta\tilde{v}_{g_{aniso}}$ can be reduced by increasing N_{λ_0}. Furthermore, it can be observed from Table 5.2 and Figure 5.5 that, for a constant grid resolution, square grid cells ($R_y = 1$) give rise to a lower anisotropy of the numerical group velocity than cells with $R_y \neq 1$. Hence, square cells are to be used. However, the simulation platform has limited memory and thus, arbitrarily large N_{λ_0} are computationally expensive. The FDTD simulator described in section 3.3 is subject to the memory requirements given by equations (3.15) and (3.19). Thus, it can be seen that $N_{\lambda_0} = 10$ requires 1.3 GB for an scenario of size $(D_x, D_y) = (77, 47)$ meters at $f_0 = 3.5$ GHz. If the grid resolution was $N_{\lambda_0} = 11$, the memory requirements would be 1.6 GB. These simulations were performed on the TESLA C870 card (see Table 3.1) which has 1.5 GB of global memory. Thus, $N_{\lambda_0} = 10$ corresponds to the largest spatial resolution that can be simulated on such a platform for this scenario. Hence with these simulation parameters ($R_y = 1$ and $N_{\lambda_0} = 10$) and as seen from Table 5.2, the anisotropy error of the numerical group velocity is $\Delta\tilde{v}_{g_{aniso}} \approx 2.65\%$. Further, this corresponds to the maximum relative error in the prediction of the time of arrival of multipath components in CIR simulations. To better comprehend the significance of this error in relationship with the results from FDTD simulations let us suppose, for instance, that two multipath components (MCs) are predicted to reach a given location at times τ_1 and τ_2. Consider also that physical propagation is exempt of anisotropy of the group velocity. Hence, it can be said that the time separation $\tau_2 - \tau_1$ between those MCs mirrors the reality within a maximum error of $\Delta\tilde{v}_{g_{aniso}}$.

What about the accuracy with which the time of arrivals (ToAs) are predicted by the simulation? In other words, how sure can we be that an MC predicted to

arrive at time τ_1 will actually arrive at that time in the physical world? To answer this question and in a similar manner to the physical phase-velocity error described by equation (4.4), the physical group velocity error can also be defined as

$$\Delta\tilde{v}_{g_{physical}} = \frac{max\{|\tilde{v}_g - v_g|\}}{v_g}. \tag{6.1}$$

It is this variable that quantifies the fidelity in the prediction of ToAs. Furthermore, a quantity called the group index that relates the group velocity to c can also be defined as

$$n_g = \frac{c}{v_g}. \tag{6.2}$$

In addition and since $\tilde{v}_g$ varies for different propagation directions, the numerical group index $\tilde{n}_g$ is also anisotropic and the simulation settings must be such that the average numerical group velocity $(\tilde{v}_g)_{avg}$ equals the physical group velocity in order to reduce $\Delta\tilde{v}_{g_{physical}}$. Further, the group index can also be expressed in terms of the refractive index as

$$n_g = n - \lambda\frac{\partial n}{\partial\lambda}. \tag{6.3}$$

However, at the radio frequencies under consideration (non-optical frequencies), n is approximately constant (see Table 4.2) in common construction materials. Thus $\partial n/\partial\lambda \approx 0$, $n_g \approx n$ and $v_g \approx v_p$. Therefore, in order to minimise $\Delta\tilde{v}_{g_{physical}}$, the properties of the numerical materials involved in the simulation must be such that the average numerical group velocity equals the physical phase velocity. Following then the same approach as in section 4.2 for the minimisation of the physical phase velocity error, the material properties to use are

$$\tilde{\mu}_r = 1 \tag{6.4}$$

$$\tilde{\varepsilon}_r = \left(\frac{(\tilde{v}_g)_{avg}}{v_p}\right)^2. \tag{6.5}$$

This results in the material parameters presented on Table 6.1, which will then be used in the simulations to be performed in this section. In addition, equations (5.27), (5.21) and (5.44) revealed that the minimum, maximum and thus also the average value of $\tilde{v}_g$ depend mainly on N_{λ_0} and S. Hence, as long as N_{λ_0} and S remain constant, the same $\tilde{\mu}_r$ and $\tilde{\varepsilon}_r$ values can be used for CIR simulations at different carrier frequencies. Finally and in order to reduce the duration of the simulations, the Courant number is chosen close to the stability limit as $S = 0.95$.

Table 6.1: Material parameters for 2D CIR simulation with $N_{\lambda_0} = 10$ and $S = 0.95$ to mimic the behaviour of the materials described in Table 4.2

Material	$\tilde{\varepsilon}_r$	$\tilde{\mu}_r$
Air	0.970	1
Plaster	2.424	1
Glass	6.185	1
Wood	2.007	1
Concrete	2.075	1

6.1.2 Control over the spread of numerical pulses

According to Goldsmith [63],

> We say that two multipath components with delay τ_1 and τ_2 are *resolvable* if their delay difference significantly exceeds the inverse signal bandwidth.

This is equivalent to saying that two multipath pulses from a CIR prediction are resolvable if their delay difference significantly exceeds their duration. Since the duration of pulses increases due to numerical dispersion, care must be put in the source design to avoid numerical pulses overlapping each other.

MCs that contribute significantly to the CIR power content have usually powers within a dynamic range of 30 dB (see for instance [64] or [22]). Simulations have shown that such multipath components reach the receiving locations within 200 ns from emission in indoor-to-outdoor scenarios, and much sooner in indoors. Other MCs arriving later suffer more attenuation and do not contribute significantly to the power delay profile. At the speed of light in a vacuum, the distance covered by these pulses is approximately 60 meters. Therefore, the Gaussian source is designed according to the criteria introduced in section 5.3 such that the duration of the received pulses after propagating $d = 60$ meters is as low as possible. Using then the results from section 5.3, the optimum source duration is $t_w = \sqrt{m_T \cdot d}$, with m_T the slope of the linear approximation to the numerical phase delay per unit of distance.

For a constant grid resolution, m_T varies between different carrier frequencies. Therefore, the received pulses also have different durations for different values of f_0. Furthermore, lower carrier frequencies suffer larger pulse spreads due to larger

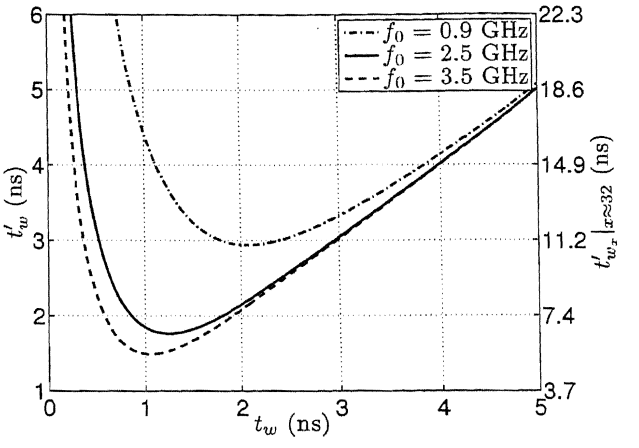


Figure 6.1: Maximum pulse duration after propagating $d = 60$ meters in a 2D staggered Yee grid with the simulation parameters of Table 6.2.

spatial steps. Hence, the source parameter t_w to be used in the wideband simulation at the spectral bands under study is selected in terms of the simulation with a lower carrier frequency. From Figure 6.1 it can be seen that the optimum source at $f_0 = 0.9$ GHz has $t_w \approx 2$ ns. This source produces a maximum pulse duration of $t'_{w_x} \approx 10.9$ ns when measured 30 dB below the power maximum ($x = 10^{30/20} \approx 32$) and after propagating $d = 60$ meters. For the remaining carrier frequencies to be simulated, the pulse spread is lower. Since the time span of 200 ns for the arrival of significant MCs is much larger, 10.9 ns is sufficient for the resolution of multipath components. Then, all simulated sources have been designed with $t_w = 2$ ns for consistency.

Regarding the conductivity values to be used in the simulations, those resulting from the calibration process at FRF = 1 in chapter 4 for each of the carrier frequencies have been used. Nevertheless, a calibration procedure like the one described in [64], in which the calibration is performed in terms of the power of the multipath components, could also be applied.

6.2 Sites

In the following, the simulation areas are described.

Table 6.2: Simulation settings for 2D CIR simulations.

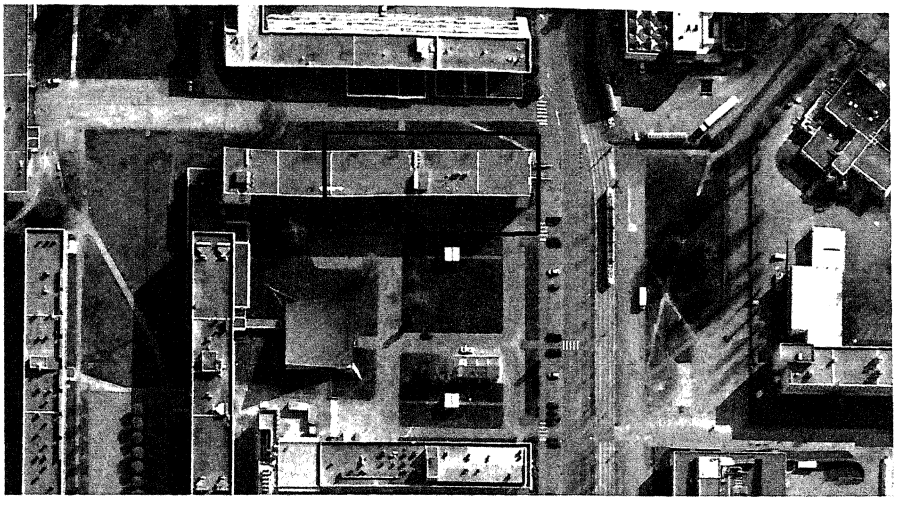
Parameter	Value
N_{λ_0}	10
S	0.95
R_y	1
t_w	2 ns
f_0	0.9, 2.5, 3.5 GHz

6.2.1 Indoor site:

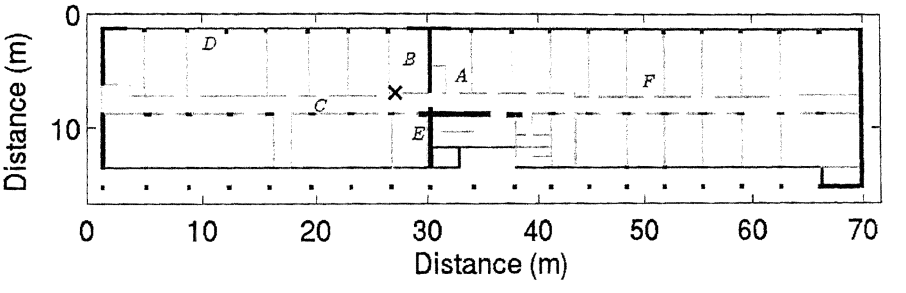
To illustrate the accuracy with which MCs are identified using the current wideband simulation method, an indoor CIR prediction has been performed. This has been done at a carrier frequency of $f_0 = 2.36$ GHz and the indoor site is a floor of the Centre of Innovation in Telecommunications and Integration of services (CITI) at the Institut National des Sciences Appliquées (INSA) in Lyon (France). This scenario is illustrated on Figure 6.2 and has been chosen due to the availability of CIR measurements data and to the existing research collaboration between CWiND and the Smart Wireless Networking (SWING) research group at CITI. In the following, this simulation site is referred to as *Site 0*. The transmitter is located in the approximate centre of the floor and the measurements have been recorded in a variety of rooms as well as the corridor.

6.2.2 Indoor-to-outdoor sites:

For the sake of generality in the wideband characterisation of femtocells, CIR predictions have been also performed in two residential femtocell sites. These are the same sites used in chapter 4 and are illustrated here on Figure 6.3 together with the mobility routes used for CIR prediction. In addition, Figure 6.4 shows a street view photograph of one of the measurement sites. In both cases, the transmitter has been located in the outermost room to resemble the case of maximum outdoor coverage. Thus, the number of walls N_w between the transmitter and the street is $N_w^{tx} = 1$. As previously mentioned, a total of three different carrier frequencies f_0 are to be analysed (see Table 6.2). Also, following the same naming scheme of chapter 4, these simulation sites are referred to as *Site 1* and *Site 2*.



(a) Aerial view of the CITI building and its surroundings.

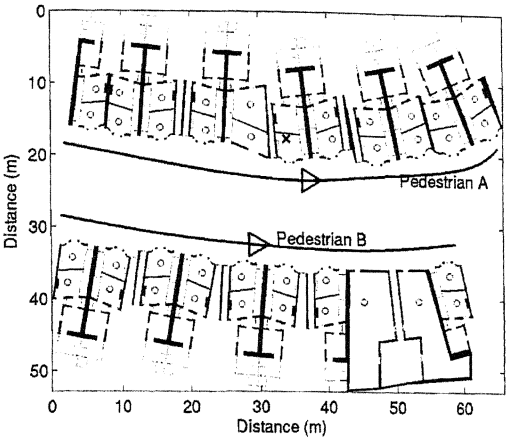


(b) Indoor scenario. The transmitter location is marked with a dark cross, and the measurement locations are indicated with capital letters.

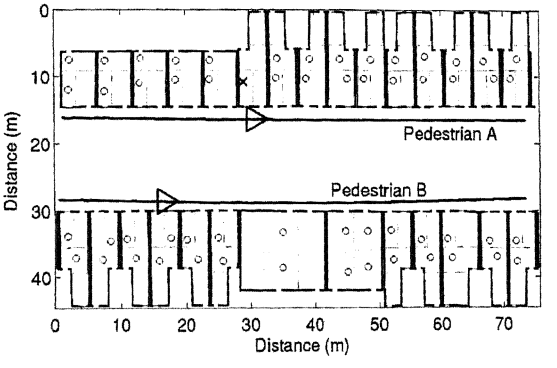
Figure 6.2: Indoor site

6.2.3 Mobility model

The following mobility model is intended to emulate the movement of pedestrians. It has been used to design the routes followed by outdoor mobile receivers, thus defining the locations where the CIR is to be recorded during simulations in the indoor-to-outdoor sites. Then, in order to simulate a time-varying channel, the receiver is modelled such that it moves at pedestrian pace along an outdoor route in the indoor-to-outdoor simulation sites described above. The average speed of the pedestrians is $\bar{v} = 1.5$ m/s and the speed standard deviation is $\sigma_v = 0.2$ m/s [65]. The routes are then designed according to the criteria described in the following paragraphs.



(a) Site 1.



(b) Site 2.

Figure 6.3: Indoor-to-outdoor sites, pedestrian routes with walking direction and indoor sounding points.

First of all, several waypoints (typically four or five) are manually selected along the street sidewalk. These include the starting and finishing route points. Then, starting from the first one, the pedestrian moves sequentially from the current waypoint towards the next one in a cubic spline trajectory. However, the direction of movement is noisy and it changes randomly according to a Poisson distribution in which the average time between direction changes is $\mu_d = 2$ seconds. The new direction is then modelled as an angular deviation from the original trajectory and following a Gaussian distribution $\mathcal{N}(0, \pi/10)$ in radians. When the pedestrian realises that it is going in the wrong direction, it self-corrects the trajectory towards the next waypoint. This occurs after a reaction time of $RT = 200$ ms.

Figure 6.4: Street-level view of site 1.

Table 6.3: Parameters of mobility model

Parameter	Value
$\bar{v}$	1.5 m/s
σ_v	0.2 m/s
μ_s	4 s
μ_d	2 s
a_M	$0.571 \cdot g$
RT	200 ms
Δt	20 ms

Furthermore, the pedestrian does not walk at a constant speed. The speed varies according to a Poisson distribution with an average time between speed changes of $\mu_s = 4$ seconds. The new speed is then chosen randomly from a Gaussian distribution $\mathcal{N}(\bar{v}, \sigma_v^2)$. In addition, the changes in speed do not happen stepwise, but gradually. To do that, an acceleration/deceleration rate is chosen from a uniform distribution $\mathcal{U}(0, a_M)$ with $a_M = 0.571 \cdot g$ and $g = 9.81\ m/s^2$ the acceleration of gravity on Earth. Such an approach is inspired by [66] and all parameters are summarised on Table 6.3.

Finally, the sampling period of the mobility route is $\Delta t = 20$ ms. According to the Nyquist sampling theorem, the maximum Doppler frequency that can be modelled with such a temporal step is $f_{D,max} \leq 1/(2 \cdot \Delta t) = 25$ Hz. At the highest

frequency to be analysed (f_0 = 3.5 GHz), such a Doppler shift is only achieved for speeds of $v_{max} = \lambda_0 * f_{D,max}$ = 2.14 m/s, which is equivalent to walking fast. However, since only pedestrian routes are being considered, Δt = 20 ms is enough for a proper characterisation of the Doppler spectrum.

6.3 Simulation results

The channel impulse response predictions and wideband channel parameters obtained with the FDTD setup explained in the previous sections are presented next.

6.3.1 Indoor site

Figure 6.5 presents the CIR measurements made at CITI in comparison with the resulting FDTD predictions. The power levels have been normalised to 0 dBm for an easier visual inspection. It must be pointed out that the empirical measurements are subject to a high noise level (the noise floor is easily identified in Figure 6.5), while the FDTD simulations are noise free. This highlights the fact that low-power MCs are more easily identifiable from numerical computations, as the noise floor does not impose any limit to their detection.

Multipath components are usually the result of large groups of reflectors and scatterers called clusters. This can also be observed in Figure 6.5, where the CIR peaks of the empirical measurements correspond to several independent MCs in the FDTD simulation. Such a fine identification of the multipath components within each cluster could be of help, for instance, in the obtainment of the exponential decay curves of the Saleh-Valenzuela model as in [59]. However, if the identification of clusters was the main objective, simulations with longer pulses (lower bandwidth) could also be performed at a reduced computational cost by applying the theory of sections 5.2 and 5.3.

From Figure 6.5, it can also be observed that some multipath components have been measured with excess delays longer than 500 ns. These have been associated with reflections on nearby buildings (see Figure 6.2(a)), which are not modelled in the numerical domain and are hence not present in the FDTD simulation. Similarly and since the current finite-difference model is based on a 2D formulation, the reflections

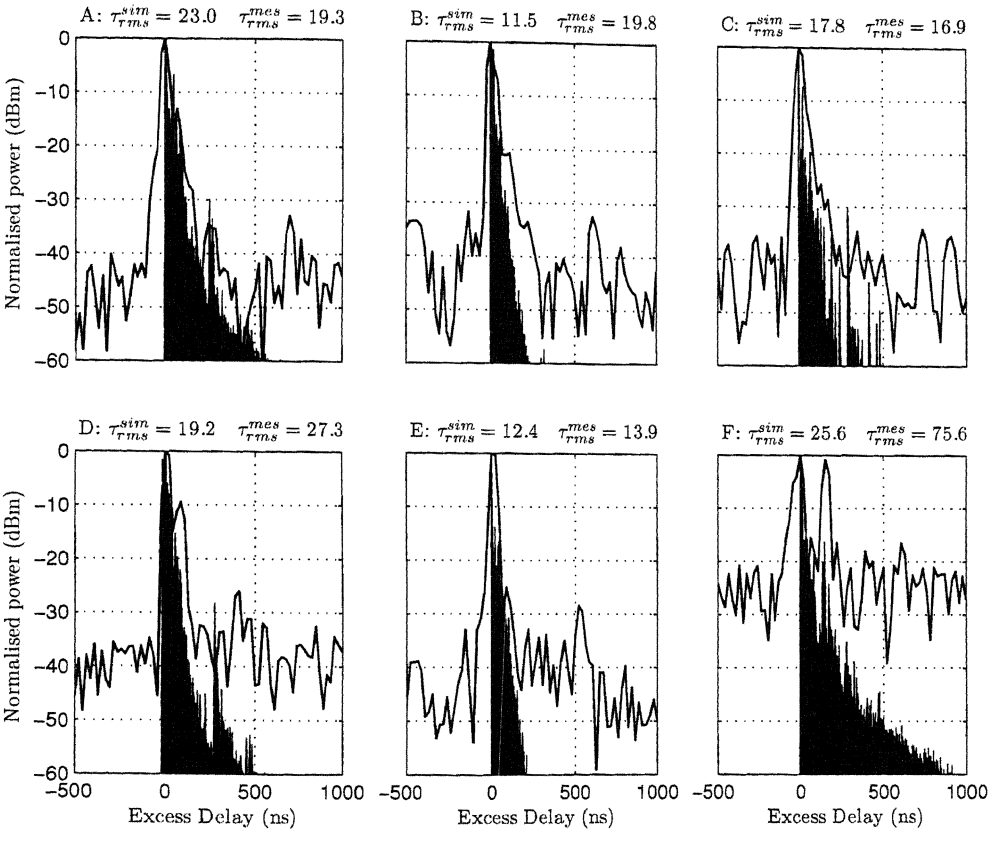


Figure 6.5: Measured CIR (solid line) and FDTD simulations (shaded area) on Site 0 (see Figure 6.2). The delay spread τ_{rms} values extracted from simulations and measurements are also indicated in ns for validation.

on floors and ceilings have not been modelled. This causes a lack of clusters in the FDTD simulation results.

In general, the ToA of the most significant MCs has been accurately predicted. However, some disagreement in the power levels can also be observed (see for example the second largest MC at location F). This is due to the use of narrowband-calibrated conductivity values in simulation, which can be adjusted through an explicit power calibration of the received pulses [64]. Nevertheless, most high-power MCs (within a range of 30 dBm) have been properly predicted both in time of arrival as well as in power.

6.3.2 Indoor-to-outdoor site

Due to the lack of physical wideband measurements on Sites 1 and 2, the results presented in this section must be handled with prudence. This is due to the current absence of wideband measurements in indoor-to-outdoor scenarios, which require complicated setups for high-power pulsed emissions. However and as will be seen, the results are coherent with the existing literature in indoor scenarios. Thus, they still serve as a reference for future empirical studies of UWB indoor-to-outdoor transmission, as well as wideband femtocell systems such as LTE-Advanced. The wideband channel parameters to be extracted from the CIR predictions are those, that are useful for system design: delay spread, coherence bandwidth, Doppler spread and coherence time. These are calculated from the numerical channel functions of the following propagation scenarios in both sites:

- Outdoor pedestrian walking along the sidewalk next to the transmitter premises. This is referred to as *Pedestrian A* on Figure 6.3.
- Outdoor pedestrian walking along the sidewalk across the street from the transmitter premises. This is referred to as *Pedestrian B* on Figure 6.3.
- Indoor static receivers. These are indicated by circles on Figure 6.3.

6.3.2.1 Channel functions

A mobile channel is fully described by its time-variant channel impulse response (CIR). However, different channel functions can also be defined to extract further

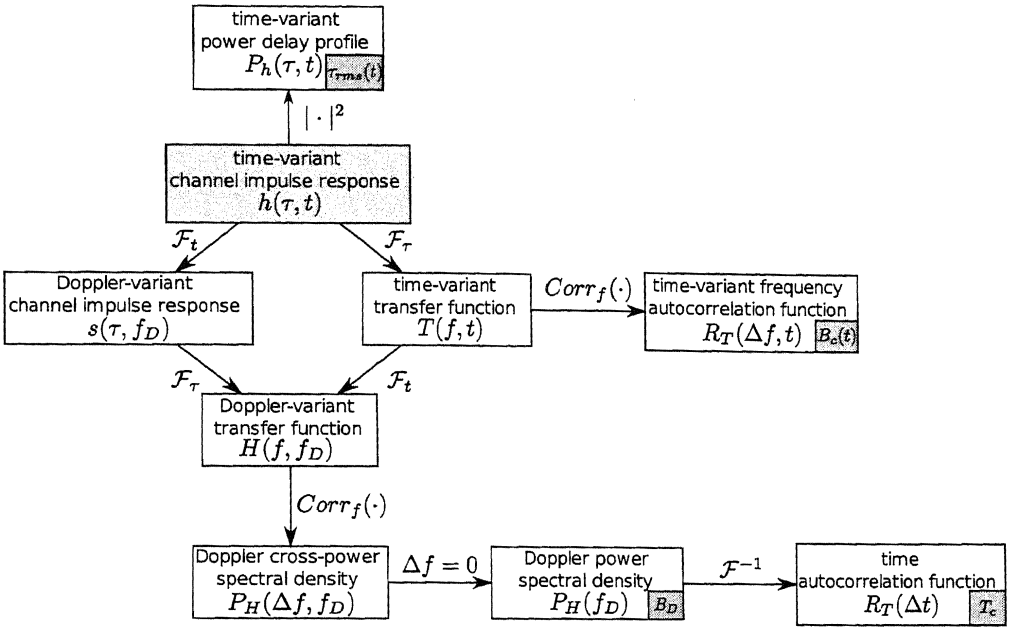


Figure 6.6: Bello functions for a deterministic channel. The time-variant channel impulse response $h(\tau,t)$ is the output electrical field of the FDTD simulation.

parameters and to analyse the channel behaviour on other domains. The channel functions were first introduced by Bello [67] for wide-sense stationary channels with uncorrelated scattering (WSSUS) and are currently widely used to describe time varying mobile channels. Following the notation of [68], the Bello functions that can be obtained from a deterministic channel model are illustrated on Figure 6.6. These relationships also indicate the procedure that must be followed to extract wideband channel parameters from deterministic CIR predictions. Since the focus is in the obtainment of four wideband parameters, simulation examples of channel functions that lead to those parameters are presented in the following.

Time-variant power delay profile (PDP): The PDP $P_h(\tau)$ provides the power of MCs associated with their respective delays. It is usually computed as the average of several measurements of the physical CIR to compensate for time-varying fading. However, in a deterministic channel model such as FDTD, no averaging is necessary. Thus, the PDP is in this case the same as the instantaneous power [69] (see Figure 6.6). Obtained in this way from the FDTD simulations, Figure 6.7 illustrates a time-

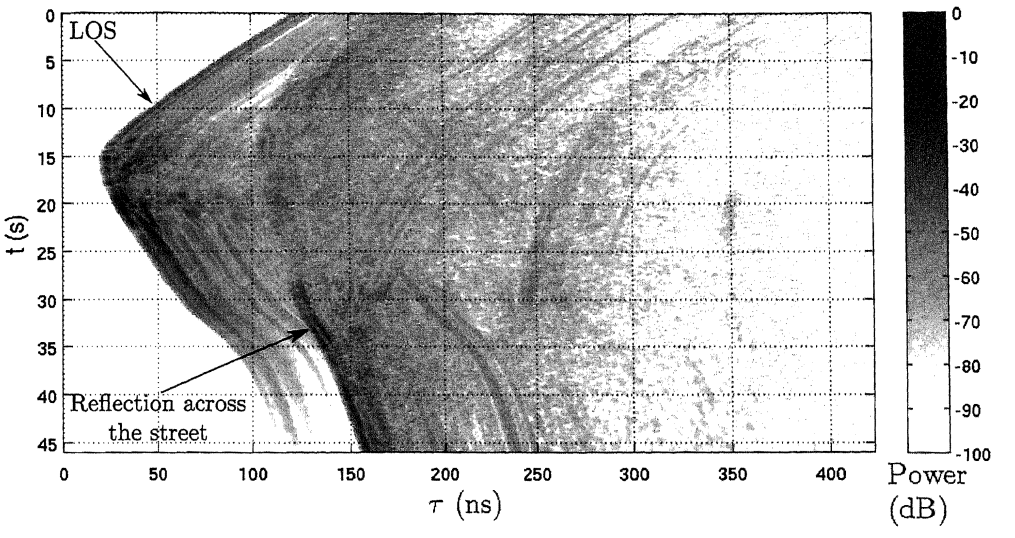


Figure 6.7: Normalised time-variant power delay profile for pedestrian A in Figure 6.3(a) at 3.5 GHz.

variant PDP for a pedestrian user. From this figure, the most meaningful clusters can be visually identified:

- In indoor-to-outdoor residential scenarios, there are two main multipath components: The line-of-sight (LOS) component, plus a reflection on the houses across the street from the transmitter premises. The LOS component corresponds, in this case, to the straight line between the receiver and the outer wall of the transmitter premises.

- In this example, at $t \approx 33$ seconds, the LOS component vanishes due to the pedestrian turning on a curve (see Figure 6.3(a)). When this happens, a decrease in the root mean square (RMS) delay spread is to be expected. There is thus a fundamental difference between the CIR in LOS and non-LOS conditions.

- The LOS component has a long tail of MCs. This is due to sequential reflections of the first pulse within the house hosting the transmitter.

Doppler power spectral density (DPSD): The DPSD $P_H(f_D)$ from one FDTD simulation is displayed on Figure 6.8. From this, it can be seen that most of the

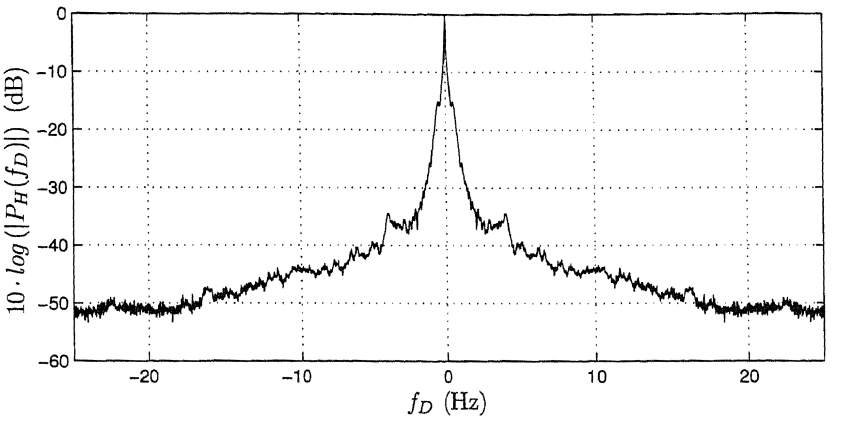


Figure 6.8: Normalised Doppler power spectral density for pedestrian A in Figure 6.3(a) at 3.5 GHz.

power is concentrated around $f_D = 0$, giving rise to a Gaussian DPSD. This spectrum is common in sites with anisotropic scattering [70], i.e. those in which the multipath components are not uniformly distributed with respect to the angle of arrival. Since in this case there is a strong LOS component, the MCs are clearly not uniformly distributed. Hence, such a DPSD could already be suspected from the time-variant PDP on Figure 6.7, in which two strong MCs could be easily identified.

Time-variant frequency autocorrelation function (FAF): This is useful to estimate the coherence bandwidth along the pedestrian route. For instance, Figure 6.9 shows the instantaneous FAF at a given time instant, which is directly mappable to a position on the route. Further, this function varies as the user moves, giving thus rise to lower coherence bandwidths in regions of more intense multipath. A recorrelation of spectral components can also be observed for certain frequency separations. However, the coherence bandwidth is defined for a contiguous set of frequencies in order to bound the areas of the spectrum in which the channel is stationary. Thus, the shaded region applies only to a contiguous frequency range.

Time autocorrelation function (TAF): This function can be calculated to assess the channel coherence time, and it is estimated as the inverse Fourier transform of the DPSD. In Figure 6.10, the time autocorrelation function (TAF) is illustrated together with the region that defines the coherence time of the channel.

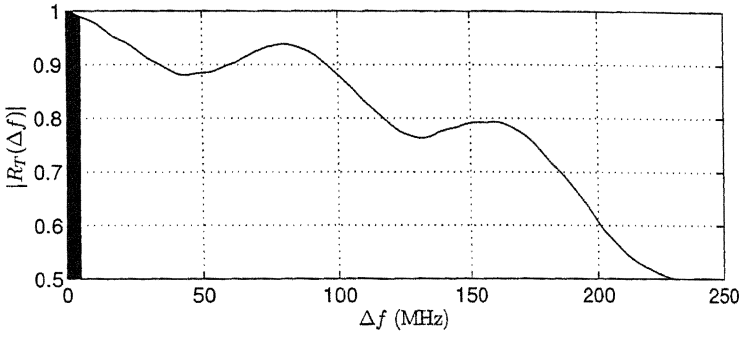


Figure 6.9: Normalised frequency autocorrelation function for pedestrian A in Figure 6.3(a) at 3.5 GHz and $t = 20$ s. The shaded area indicates the coherence bandwidth defined as the range in which $R_T(\Delta f) > 0.99$.

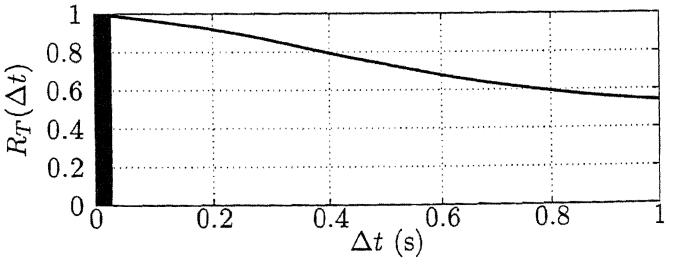


Figure 6.10: Normalised time autocorrelation function for pedestrian A in Figure 6.3(a) at 3.5 GHz. The shaded area indicates the coherence time defined as the range in which $R_T(\Delta t) > 0.99$.

6.3.2.2 Channel parameters

Next, different wideband parameters are extracted from the previously defined channel functions. This helps to illustrate a practical application of CIR simulations with FDTD: the prediction of wideband channel behaviour for the purpose of system design and calibration.

Delay spread: The RMS delay spread τ_{rms} is a measure of the multipath richness of a wireless channel and is defined as the second central moment of the power delay profile $P_h(\tau)$:

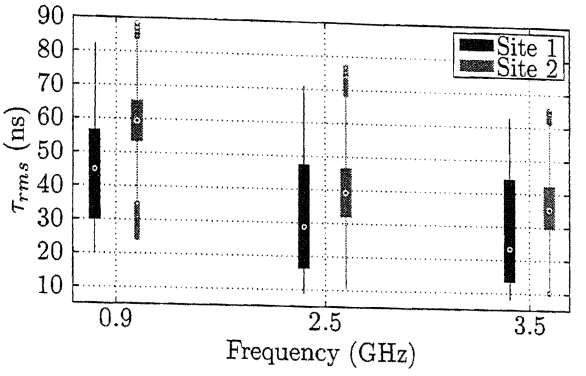
$$\bar{\tau} = \frac{\int_0^\infty \tau \cdot P_h(\tau) d\tau}{\int_0^\infty P_h(\tau) d\tau} \tag{6.6}$$

$$\tau_{rms} = \sqrt{\frac{\int_0^\infty (\tau - \bar{\tau})^2 \cdot P_h(\tau) d\tau}{\int_0^\infty P_h(\tau) d\tau}} \tag{6.7}$$

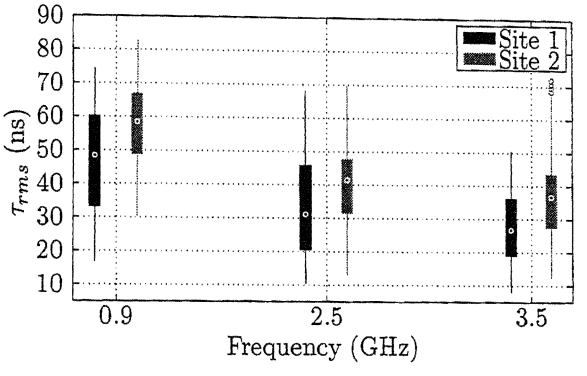
with $\bar{\tau}$ the average multipath delay.

Following this definition, τ_{rms} values have been obtained from the FDTD simulations. Figure 6.11 shows the delay spread predictions for pedestrian receivers, from which it can be seen that the delay spread is approximately the same for both pedestrians. Hence the specific location of the receiver within the street does not seem to affect greatly the value of the delay spread. In addition, an average difference of 9.3 ns in the delay spread predictions between both sites is observed due to topological differences.

Further, Figure 6.12 shows the delay spread inside of houses that are near the transmitter premises (in the same street). From these results, no significant difference has been observed in the delay spread between houses of different sides of the street. Besides, no significant difference is discerned either between inner rooms and outer rooms. However, the RMS delay spread is found to be, on average, 4.9 ns larger indoors than outdoors. This is attributed to the richer reflective environment inside of the houses, thus giving rise to more multipath. Finally, it is concluded that the RMS delay spread in femtocell links takes values as low as 10 ns and as high as 90 ns. However on average, the range $(27, 59)$ ns is representative enough for these scenarios.



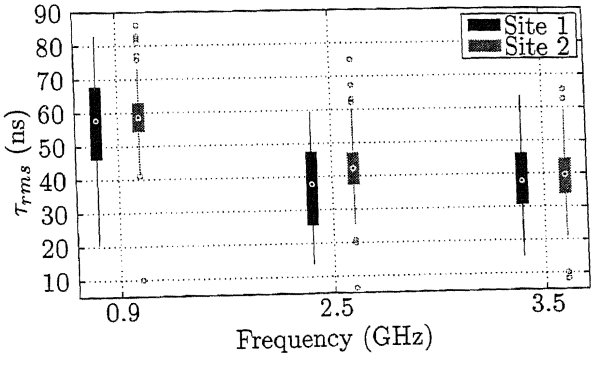
(a) Pedestrian A.



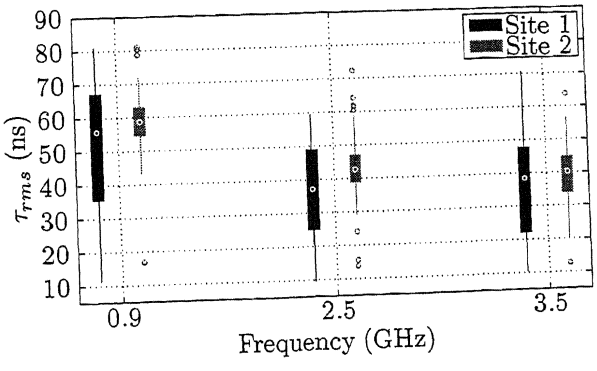
(b) Pedestrian B.

Figure 6.11: Boxplot of the delay spread due to indoor-to-outdoor propagation. The dot marks the median delay spread along a route and the box length equals the interquartile range (IQR). The whiskers extend to a maximum length of 1.5·IQR.

One final conspicuous observation of these figures is that τ_{rms} seems to decrease with frequency. However, this is identified more clearly in the decrease from 0.9 GHz to 2.5 GHz than from 2.5 GHz to 3.5 GHz. There is thus not enough evidence to identify such a trend. Still, these observations are in good agreement with the empirical indoor measurements of delay spread presented in [71], which also showed a minor decrease with frequency for NLOS conditions. Furthermore, the obtained delay spread values also fall within the range of the indoor measurements described in [72]. Finally, the ITU channel model for indoor office environments [73] also describes RMS delay spreads around 35 and 100 ns under different circumstances.



(a) Inner room ($N_w^{rx} = 2$).



(b) Outermost room ($N_w^{rx} = 1$).

Figure 6.12: Delay spread in neighboring houses.

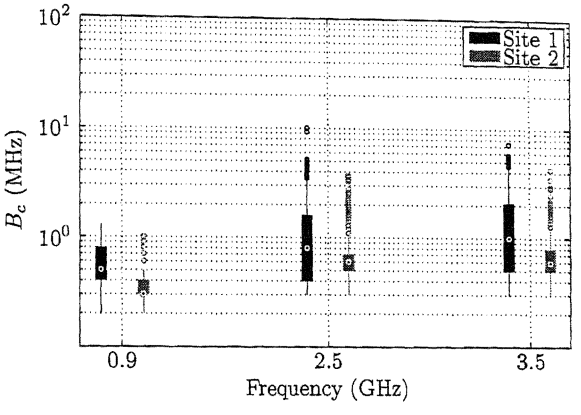
Coherence bandwidth: The coherence bandwidth B_c is a metric of the frequency separation between two correlated spectral components of a multipath channel. It can thus be defined as the frequency B_c where the FAF of the channel is above a certain threshold. Following this definition, time-variant frequency autocorrelation functions (TVFAFs) have been obtained from the numerical simulations and B_c has been calculated as the frequency separation at which the normalised TVFAF drops below 0.99. These results are illustrated on Figures 6.13 and 6.14. It must be highlighted that the coherence bandwidth and the delay spread are well known to be inversely proportional [63]. Thus, the behaviour of B_c in relation to the site topology can be explained through the previously presented results of τ_{rms}.

From the numerical simulations, the average value of the coherence bandwidth for pedestrian receivers across both sites and for all frequencies has been found to be 774 kHz. Similarly the average coherence bandwidth for indoor receivers was 602 kHz and, as seen from Figures 6.13 and 6.14, the dynamic range is approximately $[0.2, 10]$ MHz. From these results it is also seen that there is an average difference of 172 kHz between the B_c values indoors and outdoors. Finally and obviating dissimilarities in the scenarios, these results are comparable to those presented in [72] at 2.4 GHz.

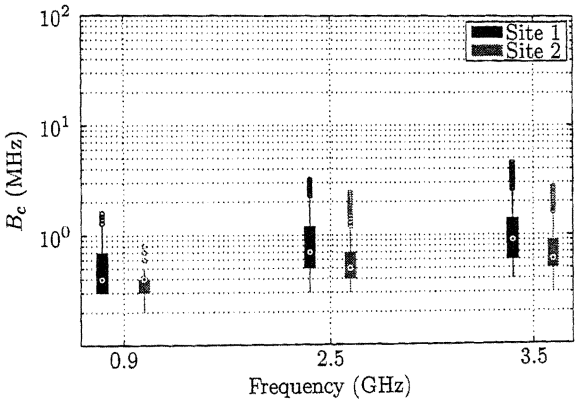
Doppler spread: When a given spectral component is transmitted through a moving multipath channel, the frequency sensed at the receiver is different from the one originally transmitted. This well-known phenomenon is the *Doppler effect*. In addition, the various angles of arrival of all received multipath components cause different Doppler shits to the same spectral component. As a consequence of this and in general, an spectrum of frequencies is received around the original frequency, thus giving rise to the DPSD previously explained. The Doppler spread B_D of a channel is a metric of the spectral broadening of the original signal and it is defined as the second central moment of the DPSD:

$$\bar{f}_D = \frac{\int_0^\infty f_D \cdot P_H(f_D) df_D}{\int_0^\infty P_H(f_D) df_D} \tag{6.8}$$

$$B_D = \sqrt{\frac{\int_0^\infty (f_D - \bar{f}_D)^2 \cdot P_H(f_D) df_D}{\int_0^\infty P_H(f_D) df_D}} \tag{6.9}$$

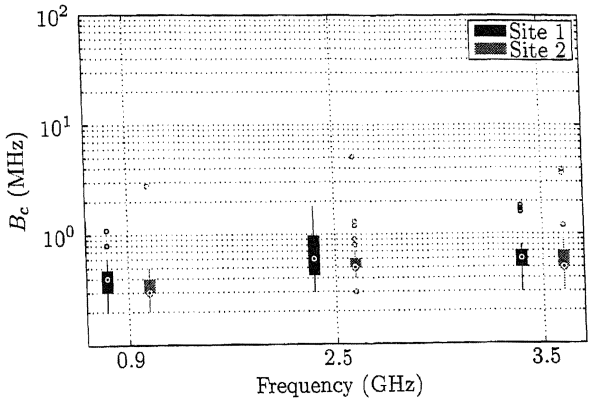


(a) Pedestrian A.

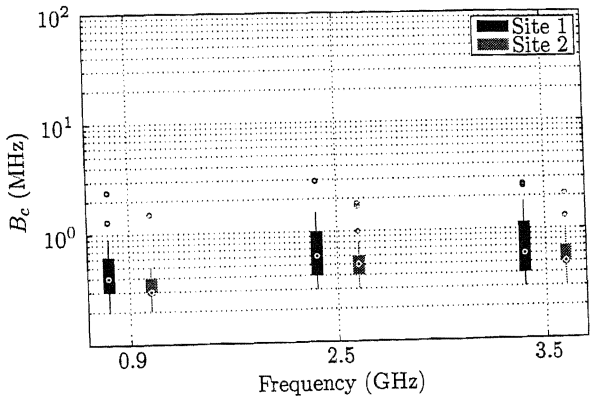


(b) Pedestrian B.

Figure 6.13: Coherence bandwidth due to indoor-to-outdoor propagation.



(a) Inner room ($N_w^{rx} = 2$).

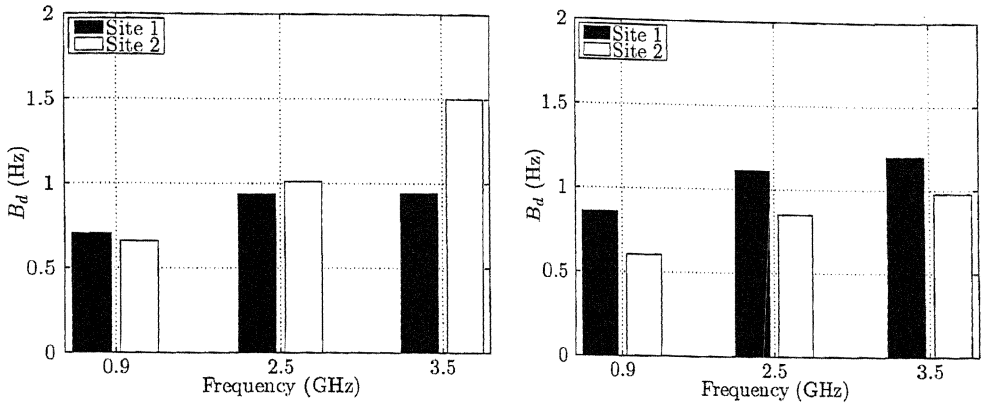


(b) Outermost room ($N_w^{rx} = 1$).

Figure 6.14: Coherence bandwidth in neighboring houses.

with f_D the Doppler frequency and $\bar{f}_D$ the average Doppler shift.

Since the FDTD model is deterministic, the CIR predictions for the indoor receivers of Figure 6.3 are static, i.e. there is no time variation. Therefore, the simulated Doppler effect is negligible for these receivers and no predictions of Doppler spread have been performed at these locations. However, applying the above definition of B_D to the simulations of pedestrian routes, Figure 6.15 shows the resulting Doppler spread values. Considering that the average speed of the pedestrians is $\bar{v} = 1.5$ m/s, it can be seen that the presented values are well below the average Doppler shift $f_{D,max} = \bar{v}/\lambda$. Further, a higher B_D is also observed at higher frequencies, which is due to a more intense Doppler effect as a result of the shorter



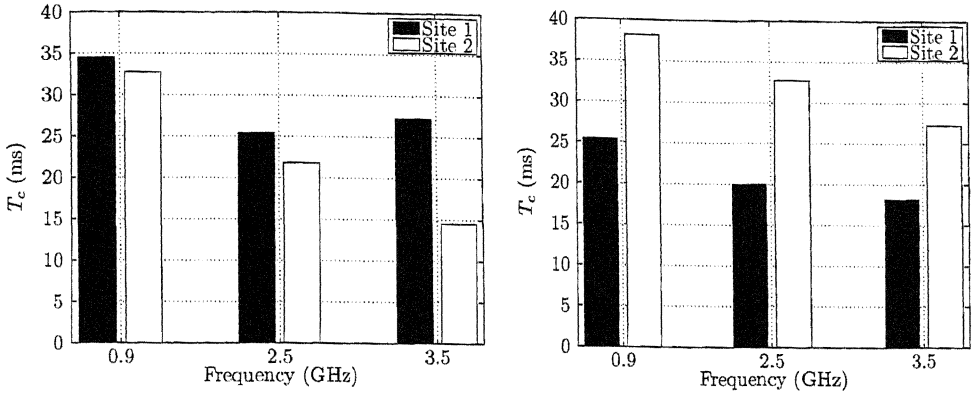
(a) Pedestrian A.

(b) Pedestrian B.

Figure 6.15: Doppler spread due to indoor-to-outdoor propagation.

wavelength. For validation, indoor measurements of Doppler spread at 910 MHz [74] have been used. These report an RMS Doppler bandwidth (i.e. Doppler spread according to equation (6.9)) of 0.87 Hz, which is akin to the outcome of the FDTD simulations.

Coherence time: The coherence time T_c is a measure of the time period during which a time-varying channel can be considered stationary, i.e. the time it takes for the channel to decorrelate. It is thus defined as the time instant T_c at which the time autocorrelation function drops below a given threshold. In FDTD simulations, the wireless channel varies from the point of view of the pedestrian receivers. Hence, T_c can also be extracted from these simulations for the purpose of channel description. To do that, T_c is calculated as the Δt at which $R_T(\Delta t)$ drops below 99% of its maximum value. The resulting T_c values are then illustrated on Figure 6.16, which are akin to those presented in [75] for the ISM band indoors. These results are useful, for instance, to balance the rate with which the channel quality indicator (CQI) is sent back to the transmitter and across the channel in femtocell links. Feedback periods below the coherence time may be unnecessary due to the channel remaining stationary. Thus, the knowledge of this channel parameter can help to increase the performance of the network.



(a) Pedestrian A.

(b) Pedestrian B.

Figure 6.16: Coherence time due to indoor-to-outdoor propagation.

6.4 Conclusions

This chapter has illustrated how the theory developed in the previous chapter can be used for the wideband characterisation of large scale radio channels. Then, results of CIR predictions have been presented in indoor and indoor-to-outdoor scenarios at different carrier frequencies. Further, bibliographic references have been given that validate the presented wideband parameters by placing them into an appropriate context.

6.4.1 Contributions

Finally, the main contribution of this chapter made by the author to the scientific community is the **wideband characterisation of the indoor-to-outdoor channel**. These results describe the behaviour of UWB systems with a transmission range beyond the transmitter premises. Furthermore, fundamental wideband channel parameters have been presented and compared to the existing literature. These are of help in the evaluation of future femtocell wideband technologies such as WiMAX or LTE.

For instance, Figure 6.11 showed that the delay spread in these scenarios is always below 90 ns. This is useful in the selection of, for instance, the cyclic prefix of OFDM systems, which must be longer than τ_{rms}. Similarly, Figure 6.13 showed that the

coherence bandwidth is higher than 200 kHz. This aids also in the design of the subcarrier spacing in OFDM systems, which must be lower than B_c to guarantee flat fading within each subcarrier. Figure 6.15 showed a maximum Doppler spread of 1.5 Hz, thus also showing how much orthogonality loss are OFDM subcarriers subject to. Finally, Figure 6.16 showed a minimum coherence time of approximately 15 ms. Considering this, the scheduler of any wireless system would only need to allocate resources to users every 15 ms, as the channel stationarity lasts always longer.

7

Concluding remarks

As a whole, this thesis pursued two global aims. The first one was to explore the potential applicability of the FDTD algorithm to the modelling of large-scale radio channels with emphasis on wideband. The second one, was to propose techniques for coverage prediction that facilitate and enhance the use of this method for the purpose of wireless networks planning and optimisation. To summarise the resulting accomplishments, the scientific contributions of this work are thus highlighted in the following section. Then, their implications are analysed and summated to situate the achievements in a broader context. Ultimately, a list of future research topics in this area is presented.

7.1 Scientific contributions

The previous chapters have presented the contributions to science after an in-depth and detailed description of the techniques involved. These achievements are abbreviated next to recapitulate and to give an encompassing view of the research undertaken:

- **Characterisation of the diffraction error due to lower-frequency FDTD simulations:** In order to save computational resources, calculations at frequencies lower than that of the physical carrier have been previously attempted. However, this thesis has shown that the attenuation due to diffractive phenomena is frequency dependent and that a measurements-based calibration

does not solve these errors. Furthermore, such an error has been quantified and an analytical formula has been devised to predict it.

- **Parallel implementation of FDTD on a graphics processing unit:** Taking advantage of the parallel nature of FDTD, this algorithm has been implemented on a GPU, thus profiting from its single instruction - multiple thread architecture. Then, this thesis has introduced and described a novel execution flow and memory management scheme for this purpose.

- **Measurements-based calibration approach of FDTD implementations using simulated annealing:** Even if a given FDTD simulator is able of running *fast*, this does not guarantee the accuracy of the results. This thesis has shown that metaheuristics can be of aid in the calibration of this model for the purpose of coverage prediction. Furthermore, this calibration approach has been validated through the evaluation of simulation results in different sites.

- **Characterisation of the path loss in indoor-to-outdoor residential scenarios:** In 2010, femtocells are commencing their worldwide deployment by part of mobile network operators. This technology has been used throughout this thesis to illustrate the applicability of FDTD to the performance of coverage predictions. Thus, the calibration and coverage results presented here help to comprehend the extent to which femtocells interfere to each other or to overlaying macrocells.

- **Optimisation of the numerical group velocity:** The speed of pulses propagating throughout the grid is determined by the numerical group velocity. Prior to the work done in this thesis, $\tilde{v}_g$ could only be approximated through numerical analysis of the dispersion relationship, which is subject to uncertainties. Now, tight bounds to $\tilde{v}_g$ have been unveiled, thus giving rise to one of the main scientific contributions of this research. In addition, this has fully characterised the anisotropy error of $\tilde{v}_g$ in simulations with pulsed sources in arbitrary media.

- **Analytical characterisation of the spread of numerical pulses:** Due to the frequency dependence of the numerical phase delay, numerical pulses in FDTD spread over the time domain. This deteriorates CIR predictions, since there is a risk that long pulses may overlap each other, thus diminishing the resolvability of multipath components. To cope with this problem, this thesis has introduced a formula that approximates the duration of Gaussian pulses after a given propagation distance and for arbitrary directions. This way, wideband FDTD simulations can be designed to assure a given degree of accuracy in the resolution of multipath components.

- **Characterisation of wideband channel parameters in indoor-to-outdoor residential scenarios:** Using the developed theory, this thesis has presented values of delay spread, coherence bandwidth, Doppler spread and coherence time at different frequencies in indoor-to-outdoor sites. These are of aid in the evaluation of wideband femtocells as well as high-power UWB systems.

7.2 The significance of this research

This section is intended to address the questions of why my work is important and where it fits in the larger field of computational electromagnetics.

Regarding the application of FDTD to the performance of CW simulations, this thesis is important for the following reasons. First of all, it has explained the consequences of using lower simulation frequencies in finite-difference models, which is an increasingly common approach within the propagation modelling community. Then, it has highlighted the parallel nature of finite-difference algorithms by showing that tremendous computational speeds can be achieved when the programming paradigm is shifted from traditional CPU-based approaches. Lastly, it has shown that empirical radio propagation measurements play a crucial role in the improvement of the prediction results, no matter how accurate the underlying site description and propagation models are.

Chapters 5 and 6 on wideband modelling holds the bulk of this work's relevance. In relation to this, I believe that this research is meaningful because of the following three reasons. First of all, it introduces the first analytical characterisation of the

propagation of numerical wavepackets in staggered Yee grids. In particular, two fundamental properties that describe the natural behaviour of FDTD and that had not been quantified before, have been presented. The second reason is that the formulation that I have introduced is generic and independent from the final application of the FDTD simulation. Because of this, the results presented here have a purpose not only for the prediction of channel impulse responses, but also for countless applications of computational electromagnetics. These include the study of photonic impulses, UWB systems, wideband antenna design, electromagnetic solitons, etc. Finally, these results explain many of the errors observed during erstwhile evaluations of wideband FDTD simulations. These minor uncertainties might have been previously disregarded by scientists and engineers using large grid discretisations. However, the mathematical tools introduced by this thesis shed light over why do these errors occur, how can they be quantified and whether or not is it worth it to relax the accuracy constraints in exchange for lower computational requirements.

7.3 Future work

Although fully coherent and practical on its own, there is still room for improvement in the techniques presented in chapters 3 and 4. To the discretion of the author, the following is a list of the more immediate research topics in this area. However, this list is not exhaustive:

1. **Effects of lower-frequency simulations on the knife-edge diffraction over multiple obstacles:** Several methods exist for the prediction of diffraction over multiple edges: Deygout, Epstein-Petersen, Giovanelli, etc. These can be compared with the results from numerical simulations for the evaluation of the simulation error in more complex scenarios.

2. **Alternative GPU-based implementations:** As this is a matter of design, this algorithm is subject to countless implementations. A promising approach lies in the use of the texture memory within the GPU for the handling of large matrices. At the time of writing this thesis, texture rendering is not supported by CUDA so this feature may be exploited using alternative platforms.

3. **Comparison of the efficiency of different calibration procedures using metaheuristics:** Although most methods reach similar solutions, some may converge faster than others. A comparative study of these would add significant value to the calibration process.

4. **Coverage prediction using 3D FDTD:** Computational advances allow today the performance of full 3D FDTD simulations at high frequencies in small indoor scenarios such as residential houses. Furthermore, larger coverage areas such as those of femtocells (the size of one street) could even be simulated on large GPU clusters. However, it is unclear whether such computationally expensive simulations increase the prediction accuracy in a significant manner with respect to calibrated 2D simulations. A proper comparative analysis of such techniques would thus be useful, specially with views to the commercialisation of GPU-based FDTD simulators.

5. **Impact of the scenario description on the accuracy of coverage predictions:** The outcome of FDTD coverage simulations depends mainly on the scenario topology. However, this can be defined roughly (e.g. assuming flat surfaces for large structures and obviating small obstacles such as paper bins), or finely (e.g. modelling minor details such as brick junctures). Thus and as suggested already in [76], it would be useful to study the impact of different degrees of scenario detail on the final coverage result.

Similarly, the following is a non-exhaustive list of research topics related to wideband FDTD simulations using finite-difference methods:

1. **Optimisation of the numerical group velocity in alternative finite-difference schemes:** The Lagrangian methodology introduced in section 5.2 has proven successful in the optimisation of $\tilde{v}_g$ for the standard FDTD. This highlighted that numerical dispersion is the main source of error in wideband simulations. However, there currently exist several low-dispersion FDTD schemes [61] that could prove more appropriate for this purpose. Nevertheless, the extrema of the numerical group velocity in such schemes is not yet known. Hence, further research on the application of the Lagrangian method to these schemes is promissing.

2. **Analytical formulae for the exact prediction of the spread of numerical pulses:** In section 5.3, an approximation to the spread of numerical pulses was introduced. This was approximated because the exact value of $\tilde{T}_p$ is the solution of a transcendental equation. Thus, the exact integrand in equation (5.66) is not analytically computable. Therefore, further research on how to model analytically the numerical phase delay for the prediction of pulse spread is necessary.

3. **Study of different calibration procedures for wideband simulations:** For this purpose, calibration targets (ToA, power of MCs, etc) must first be defined. Then, different calibration methods such as the one in [64] could be tested and compared.

4. **Impact of the scenario description on the prediction of wideband parameters:** It is well known that uncertainties in the scenario description lead to inaccuracies in the numerical simulation when compared to empirical measurements. However, the impact of these uncertainties in the extraction of wideband parameters (e.g. delay spread, ...) from pulsed FDTD simulations is unknown. Thus, an evaluation of physical scenarios with varying degrees of detail would be of interest.

References

[1] C. Sharma, "US Wireless Data Market: Q4 2009 & 2009 Update," Mar. 2010.

[2] Qualcomm, "WCDMA Network Planning and Optimisation," White Paper, May 2006. [Online]. Available: http://www.qualcomm.de/common/documents/white_papers/WCDMA_NetworkPlanningOptimization.pdf

[3] G. Mansfield, "Femtocells in the US Market - Business Drivers and Consumer Propositions," in *FemtoCells Europe 2008*. ATT, Jun. 2008.

[4] "Femtoforum," http://www.femtoforum.org.

[5] V. Chandrasekhar and J. G. Andrews, "Spectrum Allocation in Two-Tier Networks," *IEEE Transactions on Communications*, 2009. [Online]. Available: http://arxiv.org/abs/0805.1226

[6] D. Lopez-Perez, A. Valcarce, A. Ladanyi, G. De La Roche, and J. Zhang, "Intracell Handover for Interference and Handover Mitigation in OFDMA Two-Tier Macrocell-Femtocell Networks," *EURASIP Journal on Wireless Communications and Networking*, Mar. 2010.

[7] P. Kyösti *et al.*, "WINNER II Channel Models," WINER II Public Deliverable, Sep. 2007.

[8] "Predicting coverage and interference involving the indoor-outdoor interface," Ofcom, Project SES-2005-08, Tech. Rep., Jan. 2007.

[9] J. C. Maxwell, "On Physical Lines of Force," *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science*, Mar. 1861.

[10] C. Darwin, *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life.* John Murray, Nov. 1859.

[11] K. Yee, "Numerical Solution of Initial Boundary Value Problems Involving Maxwell's Equations in Isotropic Media," *IEEE Transactions on Antennas and Propagation*, vol. 14, pp. 302–307, May 1966.

[12] A. Taflove, "Application of the Finite-Difference Time-Domain Method to Sinusoidal Steady-State Electromagnetic-Penetration Problems," *IEEE Transactions on Electromagnetic Compatibility*, vol. 22, pp. 191–201, August 1980.

[13] C. Aulbert, "(Geo-)Radar-FDTD," http://carsten.welcomes-you.com/radarfdtd/.

[14] M. F. Yanik, S. Fan, M. Soljačić, and J. D. Joannopoulos, "All-optical transistor action with bistable switching in a photonic crystal cross-waveguide geometry," *Optical Letters*, vol. 28, no. 24, pp. 2506–2508, 2003.

[15] Remcom, "Overcoming UWB Antenna Design Challenges With FDTD Software," www.rfglobalnet.com, Apr. 2007.

[16] Y. Liu and C. D. Sarris, "Fast time-domain microwave CAD via an FDTD technique based on an adaptively refined/moving mesh," in *European Microwave Conference*, Apr. 2006.

[17] J. J. Simpson, A. Taflove, J. A. Mix, and H. Heck, "Advances in Hyperspeed Digital Interconnects Using Electromagnetic Bandgap Technology: Measured Low-Loss 43-GHz Passband Centered at 50 GHz," in *IEEE Antennas and Propagation Society International Symposium*, 2005, pp. 26–29.

[18] Y. Wang, S. Safavi-Naeini, and S. K. Chaudhuri, "A hybrid technique based on combining ray tracing and FDTD methods for site-specific modeling of indoor radio wave propagation," *IEEE Transactions on Antennas and Propagation*, vol. 48, no. 5, pp. 743–754, May 2000.

[19] S. Reynaud, Y. Cocheril, R. Vauzelle, C. Guiffaut, and A. Reineix, "Hybrid FDTD/UTD Indoor Channel Modeling. Application to Wifi Transmission Systems," in *IEEE Vehicular Technology Conference (VTC-Fall)*, Hyatt Regency, Sep. 2006, pp. 1–5.

[20] G. Rodriguez, Y. Miyazaki, and N. Goto, "Matrix-Based FDTD Parallel Algorithm for Big Areas and Its Applications to High-Speed Wireless Communications," *IEEE Transactions on Antennas and Propagation*, vol. 54, pp. 785–796, March 2006.

[21] A. Alighanbari and C. D. Sarris, "Rigorous and Efficient Time-Domain Modeling of Electromagnetic Wave Propagation and Fading Statistics in Indoor Wireless Channels," *IEEE Transactions on Antennas and Propagation*, vol. 55, no. 8, pp. 2373–2381, Aug. 2007.

[22] A. C. M. Austin, M. J. Neve, G. B. Rowe, and R. J. Pirkl, "Modeling the Effects of Nearby Buildings on Inter-Floor Radio-Wave Propagation," *IEEE Transactions on Antennas and Propagation*, vol. 57, no. 7, pp. 2155–2161, Jul. 2009.

[23] A. Valcarce *et al.*, "Empirical Propagation Model for WiMAX at 3.5 GHz in an Urban Environment," *Microwave and Optical Technology Letters*, vol. 50, no. 2, pp. 483–487, February 2008.

[24] G. De La Roche, "Radio wave propagation simulation in multipath environments for the study of wireless networks," Ph.D. dissertation, L'Institut National des Sciences Appliquées (INSA), Lyon, France, Dec. 2007. [Online]. Available: http://tel.archives-ouvertes.fr/tel-00257259/fr/

[25] J. M. Gorce, K. Jaffres Runser, and G. De La Roche, "Deterministic Approach for Fast Simulations of Indoor Radio Wave Propagation," *IEEE Transactions on Antennas and Propagation*, vol. 55, pp. 938–942, March 2007.

[26] B. Chopard, P. O. Lüthi, and J.-F. Wagen, "Multi-cell coverage predictions: a massively parallel approach based on the ParFlow method," in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 1, Boston, MA , USA, Sep. 1998, pp. 60–64.

[27] Y. E. Lu, "Site precise radio wave propagation simulations by time domain finite difference methods," in *IEEE Vehicular Technology Conference (VTC)*. Secaucus, NJ , USA, May 1993, pp. 875–878.

[28] A. Lauer, I. Wolff, A. Bahr, J. Pamp, and J. Kunisch, "Multi-Mode FDTD Simulations of Indoor Propagation Including Antenna Properties," in *Vehicular Technology Conference (VTC)*. Chicago, USA: IEEE, 1995, pp. 454–458.

[29] H. Claussen, "Performance of Macro- and Co-channel femtocells in a hierarchical cell structure," in *The 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sep. 2007, pp. 1–5.

[30] D. M. Sullivan, *Electromagnetic Simulation Using the FDTD Method*, 1st ed. Wiley-IEEE Press, Jul. 2000.

[31] A. Taflove and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, 3rd ed. Artech House, 2005.

[32] A. P. Zhao, "Determination of the Direction That Has Maximum Phase-Velocity for the 2-D and 3-D FDTD Methods Based on Yee's Algorithm," *IEEE Microwave and Wireless Components Letters*, vol. 13, no. 6, pp. 226–228, Jun. 2003.

[33] J.-P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *Journal of Computational Physics*, vol. 114, no. 2, pp. 185–200, Oct. 1994.

[34] J. A. Roden and S. D. Gedney, "Convolutional PML (CPML): An Efficient FDTD Implementation of the CFS-PML for Arbitrary Media," *Microwave and Optical Technology Letters*, vol. 27, pp. 334–339, June 2000.

[35] W. C. Chew and W. H. Weedon, "A 3D perfectly matched medium from modified maxwells equations with stretched coordinates," *Microwave and Optical Technology Letters*, vol. 7, no. 13, pp. 599–604, Sep. 1994.

[36] W. Yu, R. Mittra, T. Su, Y. Liu, and X. Yang, *Parallel Finite-Difference Time-Domain Method*. Artech House, 2006.

[37] R. N. Schneider, M. Okoniewski, and L. E. Turner, "A software-coupled 2D FDTD hardware accelerator," in *IEEE Antennas and Propagation Sociey International Microwave Symposium*, vol. 2, Jun. 2004, pp. 1692–1695.

[38] G. S. Baron, E. Fiume, and C. D. Sarris, "Graphics hardware accelerated multiresolution time-domain technique: development, evaluation and applications," *IET Microwaves, Antennas & Propagation*, vol. 2, no. 3, pp. 288–301, Apr. 2008.

[39] P. Sypek, A. Dziekonski, and M. Mrozowski, "How to Render FDTD Computations More Effective Using a Graphics Accelerator," *IEEE Transactions on Magnetics*, vol. 45, no. 3, pp. 1324–1327, Mar. 2009.

[40] *CUDA Programming Guide 2.3.1*, NVIDIA Corporation, Aug. 2009.

[41] "Cobham," http://www.european-antennas.co.uk/, Cobham Antenna Systems, Feb. 2010.

[42] Anritsu, http://www.eu.anritsu.com/, Feb. 2010.

[43] R. D. Yates and D. J. Goodman, *Probability and Stochastic Processes: A Friendly Introduction for Electrical and Computer Engineers*, 2nd ed. Wiley, Jun. 2004.

[44] R. Mathar, M. Reyer, and M. Schmeink, "A Cube Oriented Ray Launching Algorithm for 3D Urban Field Strength Prediction," in *IEEE International Conference on Communications (ICC)*, Jun. 2007, pp. 5034–5039.

[45] L. Jingming, J.-F. Wagen, and E. Lachat, "Effects of large grid size discretization on coverage prediction using the ParFlow method," in *International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, vol. 2. IEEE, September 1998, pp. 879–883.

[46] G. De La Roche, K. Jaffres Runser, and J. M. Gorce, "On predicting Indoor WLAN coverage with a fast discrete approach," *International Journal of Mobile Network Design and Innovation*, vol. 2, no. 1, pp. 3–12, 2007.

[47] J. M. Gorce, K. Jaffres Runser, and G. de la Roche, "FDTD based efficient 2D simulations of Indoor propagation for wireless LAN," in *IMACS, World Congress Scientific Computation, Applied Mathematics and Simulation*, Paris, France, July 2005.

[48] I. T. Union, "Recommendation P.526-10 on Propagation by Diffraction," ITU-R, 2007.

[49] W. Yu, "GEMS EM Simulation Software and Systems," www.2comu.com, Sep. 2009.

[50] "OptiFDTD: Finite-Difference Time-Domain Simulation Design Software," Optiwave, Jan. 2009.

[51] W. Lee and Y. Yeh, "On the estimation of the second-order statistics of log normal fading in mobile radio environment," *IEEE Transactions on Communications*, vol. 22, no. 6, pp. 869–873, Jun. 1974.

[52] B. D. Cullity and C. D. Graham, *Introduction to Magnetic Materials*, 2nd ed. Wiley, 2008, p. 16.

[53] A. Safaai-Jazi, S. M. Riad, A. Muqaibel, and A. Bayram, "Ultra-wideband Propagation Measurements and Channel Modeling," DARPA NETEX Program, Blacksburg, Virginia, U.S.A., Tech. Rep., Nov. 2002.

[54] L. G. Hector and H. L. Schultz, "The Dielectric Constant of Air at Radiofrequencies," *Journal of Applied Physics*, vol. 7, no. 4, pp. 133–136, 1936. [Online]. Available: http://link.aip.org/link/?JAP/7/133/1

[55] A. J. B. de Oliveira, M. T. de Melo, L. L. D. Cabral, and S. R. O. de Souza, "Plaster Glue Complex Permittivity Response in the Microwave Range," *Materials Research*, vol. 7, no. 4, pp. 565–568, 2004.

[56] A. Taflove and S. C. Hagness, *Computational Electrodynamics: the finite-difference time-domain method*, 3rd ed. Artech House, 2005, ch. 4.9.1, pp. 138–139.

[57] M. Premaratne and S. K. Halgamuge, "Rigorous Analysis of Numerical Phase and Group Velocity Bounds in Yee's FDTD Grid," *IEEE Microwave and wireless components letters*, vol. 17, no. 8, pp. 556–558, Aug. 2007.

[58] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by Simulated Annealing," *Science*, vol. 220, no. 4598, pp. 671–680, May 1983.

[59] A. Alighanbari and C. D. Sarris, "Parallel Time-Domain Full-Wave Analysis and System-Level Modeling of Ultrawideband Indoor Communication Systems," *IEEE Transactions on Antennas and Propagation*, vol. 57, no. 1, pp. 231–240, Jan. 2009.

[60] Y. Zhao, Y. Hao, A. Alomainy, and C. Parini, "UWB On-Body Radio Channel Modelling Using Ray Theory and Sub-band FDTD Method ," *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 4, pp. 1827–1835, Jun. 2006.

[61] K. L. Shlager and J. B. Schneider, "Comparison of the dispersion properties of several low-dispersion finite-difference time-domain algorithms," *IEEE Transactions on Antennas and Propagation*, vol. 51, no. 3, pp. 642–653, Mar. 2003.

[62] L. Cohen, *Time Frequency Analysis: Theory and Applications* . Prentice Hall, 1995.

[63] A. Goldsmith, *Wireless Communications.* New York, NY, USA: Cambridge University Press, 2005.

[64] J. Jemai, T. Kürner *et al.*, "Towards a Performance Boundary in Calibrating Indoor Ray Tracing Models," *EURASIP Journal on Wireless Communications and Networking*, Mar. 2009, article ID 532474. [Online]. Available: http://www.hindawi.com/journals/wcn/2009/532474.html

[65] N. Waldau, P. Gattermann, H. Knoflacher, and M. Schreckenberg, *Pedestrian and Evacuation Dynamics 2005.* Springer Berlin Heidelberg, May 2007, pp. 167–181.

[66] C. Bettstetter, "Smooth is Better than Sharp: A Random Mobility Model for Simulation of Wireless Networks," in *International Workshop on Modeling Analysis and Simulation of Wireless and Mobile Systems*, Rome, Italy, 2001.

[67] P. Bello, "Characterization of Randomly Time-Variant Linear Channels," *IEEE Transactions on Communications Systems*, vol. 11, no. 4, pp. 360–393, Dec. 1963.

[68] A. U. Sheikh, *Wireless communications: theory and techniques.* Springer, Nov. 2003.

[69] R. Vaughan and J. B. Andersen, *Channels, Propagation and Antennas for Mobile Communications*, ser. Electromagnetic Wave. The IET, 2003.

[70] "Digital mobile radio towards future generation systems," European Comission: European cooperation on the field of scientific and technical research (EURO-COST 231), 1999. [Online]. Available: http://www.lx.it.pt/cost231/

[71] T. Jamsa, V. Hovinen, A. Karjalainen, and J. Iinatti, "Frequency Dependency of Delay Spread and Path Loss in Indoor Ultra-Wideband Channels," in *The Institution of Engineering and Technology Seminar on Ultra Wideband Systems, Technologies and Applications*, Apr. 2006, pp. 254–258.

[72] S.-C. Kim, H. Bertoni, and M. Stern, "Pulse propagation characteristics at 2.4 GHz inside buildings," *IEEE Transactions on Vehicular Technology*, vol. 45, no. 3, pp. 579–592, Aug. 1996.

[73] "Guidelines for evaluation of radio transmission technologies for IMT-2000," Recommendation ITU-R M. 1225, Feb. 1997.

[74] S. Howard and K. Pahlavan, "Doppler spread measurements of indoor radio channel," *Electronics Letters*, vol. 26, no. 2, pp. 107–109, Jan. 1990.

[75] V. Papamichael, "FDTD modeling and characterization of the indoor radio propagation channel in the 434 MHz ISM band," in *International Conference on Applied Electromagnetics and Communications*, Dubrovnik, Croatia, Oct. 2003, pp. 217–220.

[76] J.-F. Wagen, "Indoor Service Coverage Predictions: How Good is Good enough?" in *European Conference on Antennas and Propagation*, Apr. 2010.

Appendix A

Lagrangian of the numerical group velocity in 2D

The system that arises from equalling equations (5.22) and (5.23) to zero can be rewritten as

$$\begin{aligned} \frac{\tilde{k}^2 - \tilde{k}_x^2}{\tilde{k}^3} + \frac{\tilde{k}_x \Delta x}{\tilde{k} \cdot tan(\tilde{k}_x \Delta x)} - \frac{\Delta x}{\Delta y} \frac{\tilde{k}_x \tilde{k}_y}{\tilde{k}^3} \frac{sin(\tilde{k}_y \Delta y)}{sin(\tilde{k}_x \Delta x)} &= \frac{\eta_g}{2A} \\ \frac{\tilde{k}^2 - \tilde{k}_y^2}{\tilde{k}^3} + \frac{\tilde{k}_y \Delta y}{\tilde{k} \cdot tan(\tilde{k}_y \Delta y)} - \frac{\Delta y}{\Delta x} \frac{\tilde{k}_x \tilde{k}_y}{\tilde{k}^3} \frac{sin(\tilde{k}_x \Delta x)}{sin(\tilde{k}_y \Delta y)} &= \frac{\eta_g}{2A}. \end{aligned} \tag{A.1}$$

For simplification, the following variables are defined:

$$\begin{aligned} \sigma_x &= \Delta x \cdot \tilde{k}_x \\ \sigma_y &= \Delta y \cdot \tilde{k}_y \\ \alpha &= \Delta x \cdot \tilde{k} \\ \beta &= \Delta y \cdot \tilde{k}. \end{aligned} \tag{A.2}$$

which are all strictly positive. Furthermore, Yee grids have at least two cells per wavelength and therefore

$$|\tilde{k}_\zeta \Delta \zeta| \leq \pi \quad \forall \zeta \in \{x, y\}. \tag{A.3}$$

Hence $\sigma_x, \sigma_y \in (0, \pi]$. Then, equalling the left hand side terms in the equations of (A.1) gives rise to the following equation:

$$\Gamma = 0 \tag{A.4}$$

where $\Gamma = \Psi_1 + \Psi_2 + \Psi_3$ and the auxiliary variables are

$$\Psi_1 = \frac{\sigma_y}{tan(\sigma_y)} - \frac{\sigma_x}{tan(\sigma_x)} \tag{A.5}$$

$$\Psi_2 = \left(\frac{\sigma_x}{\alpha}\right)^2 - \frac{\sigma_x \sigma_y}{\alpha^2} \frac{sin(\sigma_x)}{sin(\sigma_y)} \tag{A.6}$$

$$\Psi_3 = \frac{\sigma_x \sigma_y}{\beta^2} \frac{sin(\sigma_y)}{sin(\sigma_x)} - \left(\frac{\sigma_y}{\beta}\right)^2. \tag{A.7}$$

It can be easily tested that $\sigma_x = \sigma_y$ is a solution of (A.4). Therefore it is necessary to test whether this is the only solution or not. This is done in the following by studying the sign of Γ:

Case $\sigma_x > \sigma_y$: Let us first analyse Ψ_1. Since $x/tan(x)$ is a monotonically decreasing function, $\sigma_x > \sigma_y$ yields $\Psi_1 > 0$.

Let us now look into Ψ_2. Equation (A.6) can be rewritten as

$$\Psi_2 = \frac{\sigma_x}{\alpha^2 sin(\sigma_y)} (\sigma_x \cdot sin(\sigma_y) - \sigma_y \cdot sin(\sigma_x)) \tag{A.8}$$

and it can then be seen that the sign of Ψ_2 only depends on the terms in parenthesis, since $\sigma_x/(\alpha^2 sin(\sigma_y))$ is strictly positive in the interval under consideration. To analyse this, consider the following function:

$$F = \frac{\sigma_x \cdot sin(\sigma_y)}{\sigma_y \cdot sin(\sigma_x)} = \frac{sinc(\sigma_y)}{sinc(\sigma_x)}. \tag{A.9}$$

Since $sinc(x)$ is monotonically decreasing in $(0, \pi]$, it is clear that $F > 1$ when $\sigma_x > \sigma_y$ and thus $\sigma_x \cdot sin(\sigma_y) > \sigma_y \cdot sin(\sigma_x)$. Hence $\Psi_2 > 0$ for $\sigma_x > \sigma_y$.

In a similar way, Ψ_3 can be rewritten as

$$\Psi_3 = \frac{\sigma_y}{\beta^2 sin(\sigma_x)} (\sigma_x \cdot sin(\sigma_y) - \sigma_y \cdot sin(\sigma_x)). \tag{A.10}$$

Using the previous result, it turns out that $\Psi_3 > 0$ when $\sigma_x > \sigma_y$.

Therefore, $\Gamma > 0$ when $\sigma_x > \sigma_y$.

Case $\sigma_x < \sigma_y$: Using the results from the previous case, it is straightforward to prove that $\Psi_1 < 0$, $\Psi_2 < 0$ and $\Psi_3 < 0$ when $\sigma_x < \sigma_y$. Therefore $\Gamma < 0$ in this case.

Case $\sigma_x = \sigma_y$: In this case, $\Gamma = 0$.

Thus, the sign of Γ is constant when $\sigma_x > \sigma_y$ and $\sigma_x < \sigma_y$. Furthermore, Γ is a continuous function in the domains of σ_x and σ_y. Therefore, $\sigma_x = \sigma_y$ is the only solution of (A.4), which yields

$$\Delta x \cdot \tilde{k}_x = \Delta y \cdot \tilde{k}_y. \tag{A.11}$$

Appendix B

Lagrangian of the numerical group velocity in 3D

By equalling to zero equations (5.18), (5.19) and (5.20), the following system of equations arises:

$$\begin{aligned}
&\frac{\tilde{k}^2-\tilde{k_x^2}}{\tilde{k}^3}+\frac{\tilde{k}_x\Delta x}{\tilde{k}\cdot tan(\tilde{k}_x\Delta x)}-\frac{\Delta x}{\Delta y}\frac{\tilde{k}_x\tilde{k}_y}{\tilde{k}^3}\frac{sin(\tilde{k}_y\Delta y)}{sin(\tilde{k}_x\Delta x)}\\
&\qquad-\frac{\Delta x}{\Delta z}\frac{\tilde{k}_x\tilde{k}_z}{\tilde{k}^3}\frac{sin(\tilde{k}_z\Delta z)}{\tilde{k}_x\Delta x}=\frac{\eta_g}{2A}\\
&\frac{\tilde{k}^2-\tilde{k_y^2}}{\tilde{k}^3}+\frac{\tilde{k}_y\Delta y}{\tilde{k}\cdot tan(\tilde{k}_y\Delta y)}-\frac{\Delta y}{\Delta x}\frac{\tilde{k}_y\tilde{k}_x}{\tilde{k}^3}\frac{sin(\tilde{k}_x\Delta x)}{sin(\tilde{k}_y\Delta y)}\\
&\qquad-\frac{\Delta y}{\Delta z}\frac{\tilde{k}_y\tilde{k}_z}{\tilde{k}^3}\frac{sin(\tilde{k}_z\Delta z)}{sin(\tilde{k}_y\Delta y)}=\frac{\eta_g}{2A}\\
&\frac{\tilde{k}^2-\tilde{k_z^2}}{\tilde{k}^3}+\frac{\tilde{k}_z\Delta z}{\tilde{k}\cdot tan(\tilde{k}_z\Delta z)}-\frac{\Delta z}{\Delta x}\frac{\tilde{k}_z\tilde{k}_x}{\tilde{k}^3}\frac{sin(\tilde{k}_x\Delta x)}{sin(\tilde{k}_z\Delta z)}\\
&\qquad-\frac{\Delta z}{\Delta y}\frac{\tilde{k}_z\tilde{k}_y}{\tilde{k}^3}\frac{sin(\tilde{k}_y\Delta y)}{sin(\tilde{k}_z\Delta z)}=\frac{\eta_g}{2A}.
\end{aligned}\tag{B.1}$$

To simplify the notation, the following variables are defined in addition to those of (A.2):

$$\begin{aligned}
\sigma_z&=\Delta z\cdot\tilde{k}_z\\
\gamma&=\Delta z\cdot\tilde{k}.
\end{aligned}\tag{B.2}$$

Then, by equalling the terms in the left hand side of (B.1) the following equations are obtained:

$$\Psi_1 + \Psi_2 + \Psi_3 + \Psi_4 = 0 \tag{B.3}$$

$$\Psi_5 + \Psi_6 + \Psi_7 + \Psi_8 = 0 \tag{B.4}$$

where the auxiliary variables in addition to those defined in Appendix A are

$$\Psi_4 = \frac{\sigma_z sin(\sigma_z)}{\gamma^2} \cdot \left(\frac{\sigma_x}{sin(\sigma_x)} - \frac{\sigma_y}{sin(\sigma_y)} \right) \tag{B.5}$$

$$\Psi_5 = \frac{\sigma_y}{tan(\sigma_y)} - \frac{\sigma_z}{tan(\sigma_z)} \tag{B.6}$$

$$\Psi_6 = \left(\frac{\sigma_z}{\gamma} \right)^2 - \frac{\sigma_y \sigma_z}{\gamma^2} \frac{sin(\sigma_z)}{sin(\sigma_y)} \tag{B.7}$$

$$\Psi_7 = \frac{\sigma_z \sigma_y}{\beta^2} \frac{sin(\sigma_y)}{sin(\sigma_z)} - \left(\frac{\sigma_y}{\beta} \right)^2 \tag{B.8}$$

$$\Psi_8 = \frac{\sigma_x sin(\sigma_x)}{\alpha^2} \cdot \left(\frac{\sigma_z}{sin(\sigma_z)} - \frac{\sigma_y}{sin(\sigma_y)} \right). \tag{B.9}$$

Applying then a similar analysis as in Appendix A, it is easy to see that $\Psi_4 = 0$ only when $\sigma_x = \sigma_y$. Furthermore, $\Psi_4 > 0$ when $\sigma_x > \sigma_y$, and $\Psi_4 < 0$ when $\sigma_x < \sigma_y$, thus denoting $\sigma_x = \sigma_y$ as the only solution of (B.3). Similarly $\Psi_8 = 0$ only if $\sigma_y = \sigma_z$, $\Psi_8 > 0$ when $\sigma_y > \sigma_z$, and $\Psi_8 < 0$ when $\sigma_y < \sigma_z$. Equations (B.6), (B.7) and (B.8) have respectively the same form as (A.5), (A.6) and (A.7) and hence their analysis is equivalent. As a consequence of this, the unique solution of (B.1) is

$$\Delta x \tilde{k}_x = \Delta y \tilde{k}_y = \Delta z \tilde{k}_z. \tag{B.10}$$

Appendix C

Monotonicity of function

Theorem 4. *The function*

$$\xi(x) = \frac{arcsin(k \cdot sin(x))}{x} \tag{C.1}$$

with $k \geq 1$ *and defined in* $x \in (0, arcsin(1/k)]$ *is a monotonically increasing function in its domain of definition.*

Proof. Let us first define

$$g(y) = arcsin(y) \tag{C.2}$$

and

$$f(y) = \frac{g(ky)}{g(y)} = \frac{arcsin(ky)}{arcsin(y)}. \tag{C.3}$$

Then $\xi(x)$ can be rewritten as

$$\xi(x) = f(y)|_{y=sin(x)}. \tag{C.4}$$

The derivative of $f(y)$ is

$$f'(y) = \frac{k \cdot g'(ky) \cdot g(y) - g(ky) \cdot g'(y)}{g(y)^2}. \tag{C.5}$$

Then, $f(y)$ is monotonically increasing only if the numerator N of $f'(y)$ is zero or positive. N can be rewritten as

$$N = k \cdot arcsin(y) \cdot \frac{1}{\sqrt{1-(ky)^2}} - arcsin(ky) \cdot \frac{1}{\sqrt{1-y^2}}. \tag{C.6}$$

Since both terms in (C.6) are positive, then N is zero or positive if the quotient Q between them is equal to or larger than one. Q can then be written as

$$Q = \frac{k \cdot arcsin(y) \cdot \sqrt{1-y^2}}{arcsin(ky) \cdot \sqrt{1-(ky)^2}}. \tag{C.7}$$

Multiplying then the numerator and denominator of the above expression by y yields

$$Q = \frac{ky}{arcsin(ky) \cdot \sqrt{1-(ky)^2}} \cdot \frac{arcsin(y) \cdot \sqrt{1-y^2}}{y}. \tag{C.8}$$

Then, Q can be expressed as

$$Q = \frac{h(y)}{h(ky)} \tag{C.9}$$

with

$$h(y) = \frac{arcsin(y) \cdot \sqrt{1-y^2}}{y}. \tag{C.10}$$

Substituting now $y = sin(x)$ into $h(y)$, it results that

$$h(x) = h(y)|_{y=sin(x)} = \frac{arcsin(sin(x)) \cdot \sqrt{1-sin^2(x)}}{sin(x)} \tag{C.11}$$

which simplifies to

$$h(x) = \frac{x}{tan(x)}. \tag{C.12}$$

The derivative of $h(x)$ is

$$h'(x) = \frac{(sin(x) \cdot cos(x)) - x}{sin^2(x)}. \tag{C.13}$$

Using now the fact that $sin(x) < x \; \forall x > 0$, the upper bound of $h'(x)$ is

$$h'(x) < \frac{(x \cdot cos(x)) - x}{sin^2(x)} = \frac{x}{sin^2(x)} \cdot (cos(x) - 1). \tag{C.14}$$

Since $cos(x) < 1$, the right hand side term in the above equation is negative. Therefore

$$h'(x) < 0. \tag{C.15}$$

Hence, $h(x)$ is an strictly decreasing function. Using this result, it can be seen from (C.9) that $Q \geq 1$ because $k \geq 1$. Thus $N \geq 0$ and also $f'(y) \geq 0$. Therefore, $\xi(x)$ is a monotonically increasing function. □

Appendix D

Numerical phase delay at low frequencies

From (5.51) it is seen that the extrema of the numerical phase delay $\tilde{T}_p$ occur along the same directions as the extrema of the numerical wave number $\tilde{k}$. Hence

$$(\tilde{T}_p)_{max} = \frac{max\{\tilde{k}\}}{\omega} \tag{D.1}$$

and

$$(\tilde{T}_p)_{min} = \frac{min\{\tilde{k}\}}{\omega}. \tag{D.2}$$

The extrema of $\tilde{k}$ are known from [32] and rewritten in equations (5.46) and (5.47). Hence $\tilde{T}_p \in [(\tilde{T}_p)_{min}, (\tilde{T}_p)_{max}]$.

Since $\tilde{T}_p(\omega)$ is undefined for $\omega = 0$, its behaviour at low frequencies is characterised by the limit of $\tilde{T}_p(\omega)$ as ω approaches zero. Let us in the following calculate this limit for (D.1):

$$\lim_{\omega \to 0}(\tilde{T}_p)_{max} = \lim_{\omega \to 0} \frac{2d}{\Delta\zeta} \cdot \frac{arcsin\left(\frac{\Delta\zeta}{c\Delta t} \cdot sin\left(\frac{\omega\Delta t}{2}\right)\right)}{\omega}. \tag{D.3}$$

Applying L'Hôpital's rule yields

$$\lim_{\omega \to 0}(\tilde{T}_p)_{max} = \frac{2d}{\Delta\zeta} \lim_{\omega \to 0} \frac{\frac{\Delta\zeta}{c\Delta t} cos\left(\frac{\omega\Delta t}{2}\right) \frac{\Delta t}{2}}{\sqrt{1 - \left(\frac{\Delta\zeta}{c\Delta t}\right)^2 sin^2\left(\frac{\omega\Delta t}{2}\right)}} = \frac{d}{c}. \tag{D.4}$$

Similarly for (D.2)

$$\begin{aligned}\lim_{\omega\to 0}(\tilde{T}_p)_{min} &= \lim_{\omega\to 0}\frac{2Sd}{c\Delta t}\cdot\frac{arcsin\left(\frac{1}{S}\cdot sin\left(\frac{\omega\Delta t}{2}\right)\right)}{\omega} \\ &= \frac{2Sd}{c\Delta t}\lim_{\omega\to 0}\frac{\frac{1}{S}cos\left(\frac{\omega\Delta t}{2}\right)\frac{\Delta t}{2}}{\sqrt{1-\frac{1}{S^2}sin^2\left(\frac{\omega\Delta t}{2}\right)}} = \frac{d}{c}.\end{aligned} \tag{D.5}$$

Since these limits are equal, it results that

$$\lim_{\omega\to 0}\tilde{T}_p = \frac{d}{c}. \tag{D.6}$$

For illustration, this behaviour of the numerical phase delay can also be verified from Figure 5.6.

Appendix E

Publications derived from this research

Book chapters:

[1] Alvaro Valcarce and Enjie Liu. Air-Interface technologies. In *Femtocells: Technologies and Deployment*, pages 69–103. Wiley, January 2010.

[2] Alvaro Valcarce and David López-Pérez. Interference in the presence of femtocells. In *Femtocells: Technologies and Deployment*, pages 145–178. Wiley, January 2010.

[3] Guillaume De La Roche and Alvaro Valcarce. Further femtocell issues. In *Femtocells: Technologies and Deployment*, pages 261–291. Wiley, January 2010.

Journal papers & letters:

[1] Alvaro Valcarce, Hui Song, and Jie Zhang. On the design of pulsed sources and spread compensation in finite-difference time-domain electromagnetic simulations. *IEEE Transactions on Microwave Theory and Techniques (to appear in).*

[2] Alvaro Valcarce and Jie Zhang. Empirical Indoor-to-Outdoor propagation model for residential areas at (0.9 - 3.5) GHz. *IEEE Antennas and Wireless Propagation Letters (to appear in).*

[3] Alvaro Valcarce, Hui Song, and Jie Zhang. Characterization of the Numerical Group Velocity in Yee's FDTD Grid. *IEEE Transactions on Antennas and Propagation (to appear in)*, 2010.

[6] David Lopez-Perez, Alvaro Valcarce, Akos Ladanyi, Guillaume De La Roche, and Jie Zhang. Intracell Handover for Interference and Handover Mitigation in OFDMA Two-Tier Macrocell-Femtocell Networks. *EURASIP Journal on Wireless Communications and Networking*, March 2010.

[5] Guillaume De La Roche, Alvaro Valcarce, David López-Pérez, and Jie Zhang. Access Control Mechanisms for Femtocells. *IEEE Communications Magazine*, January 2010.

[6] David Lopez-Perez, Alvaro Valcarce, Guillaume De La Roche, and Jie Zhang. OFDMA Femtocells: A Roadmap on Interference Avoidance. *IEEE Communications Magazine*, 47(9):41–48, September 2009.

[7] Alvaro Valcarce et al. Applying FDTD to the coverage prediction of WiMAX femtocells. *EURASIP Journal on Wireless Communications and Networking*, March 2009.

Conference papers:

[1] Alvaro Valcarce (invited speaker) and Jie Zhang. Implementing a 2D FDTD scheme with CPML on a GPU using CUDA. In *The 26th International Review of Progress in Applied Computational Electromagnetics*, Tampere, Finland, April 2010. The Applied Computational Electromagnetics Society (ACES).

[2] Alvaro Valcarce (invited speaker). Indoor-outdoor radio propagation for radio access network planning and optimisation. In *NPL Workshop on Advances, Challenges, and Opportunities in Wireless Communications*, Teddington (U.K.), September 2009. National Physical Laboratory (NPL).

[3] Alvaro Valcarce, David López-Pérez, Guillaume De La Roche, and Jie Zhang. Limited Access to OFDMA femtocells. In *IEEE PIMRC (Personal, Indoor and Mobile Radio Communications Symposium)*, Tokyo (Japan), September 2009.

[4] Alvaro Valcarce, David López-Pérez, Guillaume De La Roche, and Jie Zhang. Predicting small-scale fading distributions with Finite-Difference methods in Indoor-to-Outdoor scenarios. In *IEEE Vehicular Technology Conference (VTC)*, Barcelona (Spain), April 2009.

[5] Alvaro Valcarce, Guillaume De La Roche, and Jie Zhang. WiMAX Femtocells: Measurements and Propagation Modeling. In *COST 2100 7th Meeting*, Braunschweig (Germany), February 2009.

[6] Alvaro Valcarce, Guillaume De La Roche, and Jie Zhang. A GPU approach to FDTD for Radio Coverage Prediction. In *11th IEEE International Conference on Communication Systems*, pages 1585–1590, Guangzhou (China), November 2008.

[7] David López-Pérez, Alvaro Valcarce, Guillaume De La Roche, Enjie Liu, and Jie Zhang. Access Methods to WiMAX Femtocells: A downlink system-level case study. In *11th IEEE International Conference on Communication Systems (ICCS)*, pages 1657–1662, Guangzhou (China), November 2008.

[8] David López-Pérez, Guillaume De La Roche, Alvaro Valcarce, Alpar Jüttner, and Jie Zhang. Interference Avoidance and Dynamic Frequency Planning for WiMAX Femtocells Networks. In *11th IEEE International Conference on Communication Systems*, pages 1579–1584, Guangzhou (China), November 2008.

[9] Guillaume De La Roche, Alvaro Valcarce, David López-Pérez, Enjie Liu, and Jie Zhang. Coverage Prediction and System Level Simulation of WiMAX Femtocells. In *COST 2100 6th Meeting*, Lille (France), October 2008.

[10] Alvaro Valcarce, Guillaume De La Roche, and Jie Zhang. On the Use of a Lower Frequency in Finite-Difference Simulations for Urban Radio Coverage. In *IEEE Vehicular Technology Conference (VTC-Spring)*, pages 270–274, Singapore, May 2008.

Index